RADIO ENGINEERING
and
ELECTRONIC PHYSICS

English Edition of

РАДИОТЕХНИКА
И
ЭЛЕКТРОНИКА

Published by the American Institute of Electrical Engineers
with the aid of a grant from the National Science Foundation

No.8 August 1961

Translated and Produced by Royer and Roger, Inc.

AMERICAN INSTITUTE OF ELECTRICAL ENGINEERS

Established 1884

345 East Forty-Seventh Street
New York 17, N.Y.

Warren H. Chase, President
N.S. Hibshman, Executive Secretary
C.E. Dean, Technical Vice President, Communications
W.F. Denkhaus, Director of Publications
L.G. Abraham, Chairman, Communications Division

The English edition of RADIO ENGINEERING AND ELECTRONIC PHYSICS is published by the American Institute of Electrical Engineers with the aid of a grant from the National Science Foundation. 1961 by American Institute of Electrical Engineers. Also published under the same arrangement are the Russian electronic journals RADIO ENGINEERING and TELECOMMUNICATIONS.

RADIO ENGINEERING AND ELECTRONIC PHYSICS

(РАДИОТЕХНИКА И ЭЛЕКТРОНИКА)

Publication of the Institute of Radio Engineering and Electronic Physics,
Academy of Sciences of the USSR

Translated and Produced
by
Royer and Roger, Inc.

Translation Editor: Leonard S. Schwartz

AIEE REVIEW COMMITTEE FOR RADIO ENGINEERING AND ELECTRONIC PHYSICS
Leonard S. Schwartz
New York University College of Engineering
Chairman

A.W. Bickley
W.P. Birkemeier
T.T.W. Bucher
J.L. Callahan
G.R. Cooper
W.A. Depp
R.S. Enticknap
A. Burr Fontaine
F.E. Froelich
Paul H. Gleichauf
G.S. Glinski
Bernard Harris
R.K. Hellmann
D.E. Higginbotham
B.A. Lengyel
W. Miller
Harry Rowe Mimno
W.W. Peterson
B. Reiffen
W.G. Schmidt
Herbert Sherman
D.L. Solomon
C.A. Stutt
G.C. Sziklai
Joseph Vogelman
G.M. White
F.B. Wood
H.L. Yudkin

Subscriptions to Radio Engineering and Electronic Physics should be sent to AIEE
Special Subscription Department
41 East 28th Street, New York 16, New York

1961 Subscription rates:

	$	£
Individuals	28.50	10
Libraries, institutes, govt. agencies	57.00	20

12 issues per annum comprising approximately 1900 pages

ALLOWANCE FOR ANTENNA DIRECTIVITY PATTERN IN THE CALCULATION OF RADIO WAVE DIFFRACTION AROUND THE EARTH

N.A. Armand and B.A. Vvedenskiy

The diffraction field of a directional antenna with its plane normal to the earth's surface in investigated. It is shown that the directional properties of a low-lying antenna can be accounted for in the calculation of the diffraction field by multiplying the dipole field equations by the antenna directive gain. In the case of a high-lying antenna the diffraction field is determined by the part of the energy flux radiated from the antenna in the horizontal direction.

INTRODUCTION

In practical calculations of the diffraction of radio waves around the earth, it is customary, when a dipole source (electric or magnetic) is used, to account for the directional properties of the real antenna by simply multiplying the equations for the radiated dipole power flux by the antenna directive gain.*

It is not obvious, however, that such an operation is correct. In fact, in the penumbral zone, and particularly in the shadow zone, the character of the field is determined to a great degree by the properties of the earth's surface, and consequently the individual properties of different antennas should average out to some degree.

Any antenna can be represented as a set of dipoles. Therefore, the field of any antenna is expressed either as a sum or as an integral of the fields of these dipoles. We can confine ourselves to the case of a vertical elementary magnetic dipole, because the calculations for such a dipole can always be appreciably simplified in the UHF band, by assuming the earth's surface to be perfectly conducting.

The following suggests a simple method of calculating the diffraction field of a directional antenna, the gist of which reduces to an integration (or summation) of the known diffraction formulas for the dipole field over the totality of the dipoles comprising the antenna. We shall stop first, however, to study several formulas pertaining to the properties of a directional antenna in free space.

1. PRELIMINARY REMARK

The fields E and H of an antenna in free space (we consider a rectangular antenna with dimensions d_1 and d_2), with the magnetic moment M uniformly distributed over the aperture, obey, as is well known [1], the following formula

---

*The antenna directive gain is defined as the ratio of the power flux radiated by the antenna in the principal direction to the power flux of an isotropic radiator, the directivity coefficient being based on the radiating properties of the antenna in free space.

$$E = H = 4\cos\psi\, M\frac{e^{ikD}}{D}\,\frac{\sin\left(\frac{kd_1}{2}\sin\varphi\right)\sin\left(\frac{kd_2}{2}\sin\psi\right)}{\sin\varphi\sin\psi}, \tag{1}$$

where $k = 2\pi/\lambda$ is the wave number; D — distance from the center of the antenna to the point of reception; φ — azimuth angle; ψ — elevation angle.

In the principal direction ($\varphi = \psi = 0$) we have

$$E = H = k^2 M d_1 d_2 \frac{e^{ikD}}{D}. \tag{2}$$

The Poynting vector in this direction is

$$S = \frac{c}{8\pi}EH^* = \frac{c}{8\pi}\frac{k^4|M|^2 d_1^2 d_2^2}{D^2}. \tag{3}$$

On the other hand, in the same direction

$$S = \frac{P}{4\pi D^2}G, \tag{4}$$

where P is the radiated power and G is the antenna directive power gain relative to an isotropic radiator.

A comparison of (3) with (4) yields

$$|M| = \frac{\sqrt{\frac{2PG}{c}}}{k^2 d_1^2 d_2^2}. \tag{5}$$

2. INFLUENCE OF ANTENNA DIRECTIVITY IN THE HORIZONTAL PLANE

Following V.A. Fok [2], we write for the field of a vertical magnetic dipole

$$H_r = E_\varphi = k^2 M_0 \frac{e^{ika\theta}}{a\sqrt{\theta\sin\theta}}V(x, y_1, y_2), \tag{6}$$

where a — radius of the earth (a_{eq} — equivalent radius of the earth); θ — polar angle between the directions from the center of the earth to the transmitter and receiver (angular distance); M_0 — dipole moment; $V(x, y_1, y_2)$ — attenuation function, which is given in the shadow zone by the single term

$$V(x, y_1, y_2) = -2e^{i\pi/4}\sqrt{\pi x}\,e^{ixt}g(y_1)\,g(y_2), \tag{7}$$

where*

$$g(y) = -\frac{w_1(t-y)}{w_1'(t)}; \tag{8}$$

$$x = m\frac{D}{a_{eq}}; \qquad D = a\theta;$$
$$y = \frac{kh}{m}; \qquad m = \left(\frac{ka_{eq}}{2}\right)^{1/3}; \tag{9}$$

D — distance between the corresponding points; h_1 — height of transmitter; h_2 — height of receiver (corresponding to these heights are y_1 and y_2); $w_1(t - y)$ — Airy functions; $t = 2.338\,e^{i\pi/3}$ — the first root of the Airy function.

---

*We put $q = \infty$, which corresponds to our case.

In the case of a rectangular antenna with its plane normal to the earth's surface, we have in lieu of (6)

$$H_r = E_\varphi = \frac{k^2 M}{a} \int\limits_{-d_1/2}^{d_1/2} \frac{e^{ika\gamma}}{\sqrt{\gamma \sin \gamma}}\, d\xi \int\limits_{\eta_1}^{\eta_2} V(x, y_1, y_2)\, dh. \tag{10}$$

here η_1 and η_2 are the heights of the upper and lower edges of the antenna $(\eta_2 - \eta_1 = d_2)$; M — density of magnetic moment; γ — angular distance from the point of integration to the point of reception;

$$\cos \gamma = \sin \theta \sin \alpha \cos (\varphi - \varphi') + \cos \theta \cos \alpha; \tag{11}$$

θ — angular distance from the point of reception to the center of the antenna; α — angular distance of the integration point; φ — azimuth of the reception point; φ' — azimuth of the integration point.

Inasmuch as the angles θ, α, and γ are small, we can obtain from (11)

$$\gamma \simeq \sqrt{\theta^2 + \alpha^2 - 2\theta\alpha \cos (\varphi - \varphi')} \tag{12}$$

or for $\rho = a\gamma$, $D = a\theta$, $\xi = a\alpha$

$$\rho = \sqrt{D^2 + \xi^2 - 2D\xi \cos (\varphi - \varphi')}. \tag{13}$$

Inasmuch as $D \gg d_1$, we have

$$\rho \simeq D - \xi \cos (\varphi - \varphi'). \tag{14}$$

The first integration in (10) yields

$$H_r = E_\varphi = 2kM \frac{e^{ikD}}{D} \frac{\sin\left[\frac{kd_1 \cos (\varphi - \varphi')}{2}\right]}{\cos (\varphi - \varphi')} \int\limits_{\eta_1}^{\eta_2} V(x, y_1, y_2)\, dh. \tag{15}$$

Since the azimuth of the point of integration can be set equal to $\pi/2$, we have

$$H_r = E_\varphi = 2kM \frac{e^{ikD}}{D} \frac{\sin\left[\frac{kd_1 \sin \varphi}{2}\right]}{\sin \varphi} \int\limits_{\eta_1}^{\eta_2} V(x, y_1, y_2)\, dh. \tag{16}$$

Comparison of this formula with (1) leads to the conclusion that the directional properties of the antenna in the horizontal plane are described by one and the same function $\frac{\sin\left(\frac{kd_1 \sin \varphi}{2}\right)}{\sin \varphi}$ both in free space and in the case of diffraction.

Consequently, the directivity of antennas in the horizontal plane is maintained in the case of diffraction around the earth. In the particular case of the principal direction $(\varphi = 0)$ [see (5)] we have

$$H_r = E_\varphi = \frac{1}{d_2} \sqrt{\frac{2PG}{c}} \frac{e^{ikD}}{D} \int\limits_{\eta_1}^{\eta_2} V(x, y_1, y_2)\, dh. \tag{17}$$

We have assumed here, strictly speaking, that under free-space conditions the antenna (with the magnetic moment or field distributed uniformly over the aperture) radiates the same amount of power. This assumption is justified if the antenna is raised above the earth's surface to a height exceeding several wavelengths.* In the case of UHF antennas this condition is almost always satisfied.

---

*See, for example, [3], pp 75 — 78, where a flat earth is discussed. It is reasonable to assume that the deductions of [3] remain valid in the case of a spherical earth.

3. INFLUENCE OF ANTENNA DIRECTIVITY IN THE VERTICAL PLANE

The effect of antenna directivity in the vertical plane is more complicated. The main difficulty lies here in the fact that the attenuation function $V(x, y_1, y_2)$ cannot be integrated in general form, since the analytic form of the integral $\int_0^z w(t)\,dt$ is unknown, where $w(t)$ is the Airy function. This integration can, however, be carried out numerically, since tables of the height factors are available in [4].

We use formula (7). Then (17) is rewritten as

$$H_r = E_\varphi = -\frac{2}{d_2}\sqrt{\frac{2PG}{c}}\,\frac{e^{-ikD+\frac{i\pi}{4}+ixt}}{D}\sqrt{\pi x}\,g(y_2)\int_{\eta_1}^{\eta_2} g(y_1)\,dh \qquad (18$$

or, if we change to a new integration variable y by means of (9),

$$H_r = E_\varphi = -2\sqrt{\frac{2PG}{c}}\,\frac{e^{ikD+\frac{i\pi}{4}+ixt}}{D}\,\frac{\sqrt{\pi x}g(y_2)}{y'_2-y'_1}\int_{y'_1}^{y'_2} g(y)\,dy \quad (y'_{1,2}=k\eta_{1,2}/m). \qquad (19$$

The quantity

$$I=\frac{1}{y'_2-y'_1}\int_{y'_1}^{y'_2} g(y)\,dy, \qquad (20$$

which is equal to the height factor of the dipole averaged over the height, has the meaning of an antenna height factor.

Comparison of (20) with (7) leads to the conclusion that in our case the ratio of the anten dipole fields is not merely $\sqrt{G}$, as in free space, but $\sqrt{G}I/g(y_A)$, where y_A is the equivalent dipole height, which naturally coincides with the average equivalent antenna height $[y_A = = 1/2\,(y'_2 - y'_1)$. The appearance of the correction factor

$$L=\frac{I}{g(y_A)}=\frac{\int_{y'_1}^{y'_2} g(y)\,dy}{(y'_2-y'_1)\,g(y_A)} \qquad (21$$

is due to the diffraction.

We note that the ratio of the fields of a directional antenna and of a dipole is independent in the diffraction zone of the height of the receiving antenna. This does not take place in free space and is due to the fact that the field structure in the diffraction zone depends only o the form and electric properties of the earth's surface. Mathematically this is brought about by the fact that we have used a single-term diffraction formula for the calculations. As we move from the shadow zone to the penumbral zone, we must include more and more terms of the diffraction series in the calculation of the field intensity. Consequently, the ratio of the antenna to dipole fields will depend in the penumbral and shadow zones on the height of the reception point.

To calculate the correction factor L, we have numerically integrated the height factor g(y) with the aid of the tables of [4], and tabulated the integral

$$K(y)=\int_0^y g(t)\,dt. \qquad (22$$

The integration was by the trapezoidal formula in steps of 0.1.* The values of the integral (22) are listed in the table.

Calculations based on the tabulated values show that when the equivalent height of the antenna y'_A is changed from zero to 4.5 (the actual height changes approximately, for examp from zero to 450 meters at λ = 300 cm and from zero to 17 meters when λ = 3 cm) and when the vertical antenna dimension $y_2' - y_1'$ is changed from 0 to 1 (the actual dimension of the antenna, for example, changes from 0 to 100 meters at λ = 300 cm and from 0 to 4 meters at λ = 3 cm), the factor L differs little from unity (in our calculation the minimum value of L was 0.86).

---

*Check calculations by Simpson's rule were made by G.B. Linkovskiy, to whom the authors are sincerely grateful. These calculations agree with our own within a fraction of 1%

The result obtained signifies that, within the indicated limits of equivalent heights, the directional properties of the antennas can be accounted for, with sufficient practical accuracy, by multiplying the dipole diffraction formulas by the antenna directivity gain. Thus, the simple method of accounting for the directional properties of antennas, referred to at the start of the article, is justified for these heights. The conditions change somewhat for high antennas (see below).

4. ON THE POSSIBILITY OF INTEGRATING THE HEIGHT FACTOR

The integral (22) of the height factor can be expressed in terms of known functions if the antenna heights are relatively small. This follows from the Debye (trigonometric) approximation of the height factors, used in earlier papers on the diffraction of radio waves around the earth [5, 6, 7]. We can arrive at this approximation from our own formulas in the following manner. Since the height factor satisfies the Airy equation [8]

$$\frac{d^2g}{dy^2} + (y - t)\, g = 0 \tag{23}$$

with initial conditions

$$g(0) = 0,$$

$$g'(0) = 1, \tag{24}$$

we can, by solving equation (23) with initial conditions (24) by the WKB method (see, for example, [9, 10]), obtain

$$g(y) \simeq \frac{\sin\left[\frac{2}{3}(y - t)^{3/2} - \frac{2}{3}(-t)^{3/2}\right]}{\sqrt[4]{(-t)(y - t)}}. \tag{25}$$

It is known that this approximation is satisfactory in practice at relatively small antenna heights.

Table

Values of K(y) calculated from (22)

y	Re$K(y)$	Im$K(y)$	y	Re$K(y)$	Im$K(y)$	y	Re$K(y)$	Im$K(y)$	y	Re$K(y)$	Im$K(y)$	y	Re$K(y)$	Im$K(y)$
0,0	0,000	0,000	1,0	0,527	0,089	2,0	1,869	1,412	3,0	—0,788	5,126	4,0	—5,835	4,131
0,1	0 005	0,000	1,1	0,640	0,131	2,1	1,966	1,694	3,1	—0,322	5,460	4,1	—6,379	3,312
0,2	0 020	0,000	1,2	0,763	0,187	2,2	2,037	2,008	3,2	—0,219	5,793	4,2	—6,791	2,365
0,3	0 045	0,001	1,3	0,896	0,259	2,3	2,074	2,351	3,3	—0,829	5,945	4,3	—7,040	1,309
0,4	0,081	0,002	1,4	1,035	0,347	2,4	2,069	2,721	3,4	—1,500	6,061	4,4	—7,097	0,172
0,5	0,127	0,006	1,5	1,180	0,458	2,5	2,016	3,112	3,5	—2,220	6,072	4,5	—6,936	—1,012
0,6	0,185	0,011	1,6	1,328	0,593	2,6	1,906	3,520	3,6	—2,961	5,963	4,6	—6,540	—2,199
0,7	0,253	0,021	1,7	1,475	0,753	2,7	1,733	3,936	3,7	—3,725	5,722	4,7	—5,898	—3,343
0,8	0,333	0,036	1,8	1,617	0,942	2,8	1,492	4,350	3,8	—4,477	5,339	4,8	—5,010	—4,390
0,9	0,424	0,058	1,9	1,750	1,162	2,9	1,178	4,751	3,9	—5,190	4,808	4,9	—3,886	—5,286
1,0	0,527	0,089	2,0	1,869	1,412	3,0	0,788	5,126	4,0	—5,835	4,131	5,0	—2,549	—5,974

Note: Comma designates decimal point

In this case

$$K(y) = \int_0^y \frac{\sin\left[\frac{2}{3}(\xi - t)^{3/2} - \frac{2}{3}(-t)^{3/2}\right]}{\sqrt[4]{(-t)(\xi - t)}}\, d\xi. \tag{26}$$

If we make in this integral the substitution

$$\tau = \sqrt{\frac{2}{3}}\,(\xi - t)^{3/4}, \tag{27}$$

(26) is transformed to

$$K(y) = \frac{2\sqrt{\frac{2}{3}}}{\sqrt[4]{(-t)}} \int\limits_{\sqrt{\frac{2}{3}(-t)^{3/4}}}^{\sqrt{\frac{2}{3}(y-t)^{3/4}}} \sin\left[\tau^2 - \frac{2}{3}(-t)^{3/2}\right] d\tau. \tag{28}$$

Since the function

$$F(z) = \int\limits_0^z e^{t^2} dt \tag{29}$$

has been tabulated in the complex plane [11, 12, 13], the calculation of the integral (28) presents no difficulty. Such a calculation, however, is more cumbersome than a direct numerical integration of the height factors.

5. CASE OF LARGE HEIGHTS

The case of large heights, more accurately, the case when

$$y_A \gg 1, \tag{30}$$

calls for a special examination. The inequality (30) enables us to calculate the correction factor L with the aid of the asymptotic representation of the Airy function [8]

$$w_1(t-y) = \frac{e^{i\pi/4}}{\sqrt[4]{y-t}} e^{i\frac{2}{3}(y-t)^{3/2}} \tag{31}$$

and yields approximately (see appendix)

$$L = \frac{\sin\left[\frac{(y_2' - y_1')\sqrt{y_A}}{2}\right]}{\frac{(y_2' - y_1')\sqrt{y_A}}{2}}. \tag{32}$$

This formula has been derived subject to the satisfaction of the set of inequalities

$$d_2 \ll 2h_A, \tag{33}$$

$$h_A \gg \frac{3}{8}\sqrt[3]{3}(a_{\mathrm{eq}}\lambda^2)^{1/3}. \tag{34}$$

If we use relations (9) and express the reduced heights by the actual ones, we obtain

$$(y_2' - y_1')\sqrt{y_A} = kd_2 \sqrt{\frac{2h_A}{a_e}}, \tag{35}$$

but

$$\sqrt{\frac{2h_A}{a_{\mathrm{eq}}}} \simeq \sin\beta, \tag{36}$$

where β is the angle between the normal to the antenna plane and the horizontal plane passing through the center of the antenna. Thus, the correction factor

$$L \simeq \frac{\sin\left(\frac{kd_2 \sin\beta}{2}\right)}{\frac{kd_2 \sin\beta}{2}} \tag{37}$$

is none other than the directivity pattern of the antenna in the vertical plane, plotted for the angle β. This means that the intensity of the diffraction field is determined by that part of the field which is radiated by the antenna in the horizontal direction.

When $\sin\beta = 2\pi/kd_2 = \lambda/d_2$, according to formula (37), L should vanish, but actually more accurate calculations result in a finite, though small quantity

$$L \simeq \frac{1}{2}\frac{t}{y_A}, \tag{38}$$

(see appendix). The fact that the field falls almost to zero means that the principal lobe of the directivity pattern no longer is "tangent" to the earth's surface. The fact that a finite quantity on the order of (38) is obtained in lieu of zero is due to purely diffraction effects and is connected with the fact that when the principal lobe "breaks away" from the earth's surface, the field due to the side lobes of the directivity pattern comes into play in the diffraction zone.

CONCLUSION

1. The directional properties of an antenna in the horizontal plane are retained in the case of diffraction around the earth.

2. The height profiles of fields of antennas with different directivities are similar to each other in the diffraction zone.

The directional properties of relatively low-lying antennas can be taken into account for the diffraction field by simply multiplying the formulas for the field intensity due to the elementary dipole by the directivity of the antenna in free space.

In the case of high-lying antennas, the diffraction field of the dipole must be multiplied by the antenna directivity gain and by its directivity pattern in the vertical plane, taken for the value of the angle between the antenna axis and the horizontal plane passed through the center of the antenna.

This situation is a specific illustration of the local-field principle, according to which the field in the diffraction zone is determined by the ray tangent to the earth's surface.

We have carried out the investigations for a transmitting antenna. By the reciprocity theorem, these deductions hold also for a receiving antenna.

APPENDIX

At large values of y the asymptotic representation (31) is valid and therefore

$$L = \frac{\int\limits_{y_1'}^{y_2} g(y)\,dy}{(y_2' - y_1')\, g(y_A)} = \frac{\sqrt[4]{y_A - t}}{y_2' - y_1'}\, e^{-i\frac{2}{3}(y_A - t)^{3/2}} \int\limits_{y_1'}^{y_2} \frac{e^{i\frac{2}{3}(y-t)^{3/2}}}{\sqrt[4]{y - t}}\, dy.$$

By the substitution

$$\tau = \sqrt{\frac{2}{3}}\,(y - t)^{3/4}$$

this expression can be transformed to

$$L = \frac{2\sqrt{\frac{2}{3}}\sqrt[4]{y_A - t}\; e^{i\frac{2}{3}(y_A - t)^{3/2}}}{y_2' - y_1'} \int\limits_{\tau_1}^{\tau_2} e^{i\tau^2} d\tau,$$

where

$$\tau_{1,2} = \sqrt{\frac{2}{3}}\,(y_{1,2}' - t)^{3/4}.$$

When $|x| \gg 1$, the integral

$$\int\limits_x^\infty e^{i\xi^2} d\xi$$

has an asymptotic expansion

$$\int\limits_x^\infty e^{i\xi^2} d\xi \simeq -\frac{1}{2i}\frac{e^{ix^2}}{x} + \frac{1}{4}\frac{e^{ix^2}}{x^3} + \dots .$$

Therefore, when $|t_{1,2}| \gg 1$

$$L \simeq \frac{\sqrt[4]{y_A - t}}{i\,(y_2' - y_1')}\, e^{-i\frac{2}{3}(y_A - t)^{3/2}} \left[\frac{e^{i\frac{2}{3}(y_2' - t)^{3/2}}}{(y_2' - t)^{3/4}} - \frac{e^{i\frac{2}{3}(y_1' - t)^{3/2}}}{(y_1' - t)^{3/4}} \right].$$

Inasmuch as

$$y_{1,2} = y_A \mp \frac{y_2' - y_1'}{2},$$

we can, by putting $y_A \gg (y_2' - y_1')$ [which is tantamount to inequality (33)], obtain

$$\frac{2}{3}(y_{1,2} - t)^{3/2} \simeq \frac{2}{3}(y_A - t)^{3/2} \mp \frac{y_2' - y_1'}{2}\sqrt{y_A - t}$$

Therefore

$$L \simeq \frac{\sin\left[\frac{(y_2' - y_1')\sqrt{y_A - t}}{2}\right]}{\frac{(y_2' - y_1')\sqrt{y_A - t}}{2}}.$$

If we put furthermore $y_A \gg t$, corresponding to inequality (34), we can get

$$L \simeq \frac{\sin\frac{(y_2' - y_1')\sqrt{y_A}}{2}}{\frac{(y_2' - y_1')\sqrt{y_A}}{2}} + \frac{t}{y_A}\left[\frac{\sin\frac{(y_2' - y_1')\sqrt{y_A}}{2}}{y_2' - y_1'} - \frac{1}{2}\cos\frac{(y_2' - y_1')\sqrt{y_A}}{2}\right].$$

Discarding the second term in this formula, we obtain the needed formula (32). In particular, when $(y_2' - y_1')\sqrt{y_A} = 2\pi$ we obtain in accordance with (38)

$$L \simeq \frac{1}{2}\frac{t}{y_A}.$$

REFERENCES

1. L.A. Vaynshteyn, Elektromagnitnye volny, (Electromagnetic Waves), Izd. Sovyetskoye radio, 1957.
2. V.A. Fok, Izv. AN SSSR, Ser. fiz., 1950, 14, 1, 70.
3. B.A. Vvedenskiy, A.G. Arenberg, Voprosy rasprostraneniya ul'trakorotkikh voln, ch. I (Problems in the Propagation of UHF Waves) Part I, Izd. Sovetskoye radio, 1948.
4. P.A. Azrilyant, M.G. Velkina, Chislennye rezul'taty teorii difraktsii radiovoln vokrug zemnoy poverkhnosti, (Numerical Results of the Theory of Diffraction of Radio Waves Around the Earth's Surface), Izd. Sovetskoye radio, 1957.
5. G.N. Watson, Proc. Roy. Soc., A, 1918, 95.
6. B.A. Vvedenskiy, Zh. Tekhnich, Fiz. (USSR), 1935, 2, 624 – 639; ZhTF, 1936, 6, 163; ZhTF, 1937, 7, 1647.
7. H.B. Bremmer, Terrestrial Radio Waves, Theory of Propagation. Elsevier Publishing Company, New York, Amsterdam, London, Brussels, 1949.
8. V.A. Fok Tablitsy funktsiy Eyri, (Tables of Airy Functions), 1946.
9. Propagation of UHF Radio Waves (Translated from English) edited by V.A. Shillerov, Izd. Sovetskoye radio, 1954.
10. L.M. Brekhovskikh, Volny v sloistykh sredakh, (Waves in Layered Media), Izd. AN SSSR, 1959.
11. V.M. Fadeyeva, N.M. Terent'yev, Tablitsy znacheniy integrala veroyatnostey ot kompleksnogo argumenta, (Tables of the Values of the Probability Integral of Complex Argument), GITTL, 1954.
12. K.A. Karpov, Tablitsy funktsii $w(z) = e^{-z^2}\int_0^z e^{x^2}dx$ v kompleksnoy ploskosti, (Tables of the Function $w(z) = e^{-z^2}\int_0^z e^{x^2}dx$ in Complex Plane), Izd. AN SSSR, 1954.
13. K.A. Karpov, Tablitsy funktsii $F(z) = \int_0^z e^{x^2}dx$ v kompleksnoy ploskosti, (Tables of the Function $F(z) = \int_0^z e^{x^2}dx$ in Complex Plane), Izd. AN SSSR, 1954.

Submitted to the editors, 23 August, 1960

INVESTIGATION OF SPACE-TIME CORRELATIONS OF THE FIELD LEVEL OF CENTIMETER WAVES PROPAGATING IN A TURBULENT ATMOSPHERE

P. Ye. Kubasov

Results are given of an experimental investigation of the space-time correlation characteristics of the fluctuations of the field of a 3-cm wave propagating over the earth's surface in the direct visibility zone. Measurements of the field fluctuations at two points with variable base carried out in June – December 1959, were used to determine the normalized spatial and temporal correlation functions of the field level fluctuations; these functions were analyzed and estimates of the correlation radii obtained.

INTRODUCTION

As is well known, the level of a UHF radio signal propagating in a medium with random inhomogeneities does not remain constant either in time or in space.

Important characteristics that describe the statistical properties of such a signal are the spatial correlation function

$$f(P_1, P_2, t) = \overline{\Delta E(P_1, t)\,\Delta E(P_2, t)} = \overline{[E(P_1, t) - \overline{E(P_1, t)}] \times} \\ \overline{\times\,[E(P_2, t) - \overline{E(P_2, t)}]} \tag{1}$$

and the temporal correlation function

$$f_1(P, t, \tau) = \overline{\Delta E(P, t)\,\Delta E(P, t+\tau)} = \overline{[E(P, t) - \overline{E(P, t)}] \times} \\ \overline{\times\,[E(P, t+\tau) - \overline{E(P, t+\tau)}]} \tag{2}$$

where $E(P_1, t)$ and $E(P_2, t)$ are the field amplitudes at the points P_1 and P_2 of the region of space under consideration at the instant of time t (the superior bar denotes statistical averaging).

If the fluctuating process under consideration is stationary in time and homogeneous in space, then

$$f(P_1, P_2, t) = f(P_1, P_2), \quad f_1(P, t, \tau) = f_1(\tau).$$

Normalizing in (1) and (2) the quantities

$$\sigma^2 = f(P_1, P_2)_{P_1=P_2} \text{and} \sigma_1^2 = f_1(\tau)_{\tau=0},$$

we obtain expressions for the spatial and temporal correlation coefficients of the field amplitude fluctuations

$$\rho_E(\vec{l}) = \frac{f(P_1, P_2)}{\sigma^2}, \tag{3}$$

$$\rho_E(\tau) = \frac{f_1(\tau)}{\sigma_1^2}, \tag{4}$$

where $\vec{l}$ is the vector joining the points P_1 and P_2 in the region of space under consideration, and τ is the time shift.

The spatial and temporal correlations of the field amplitude fluctuations, defined by (3) and (4), depend, as is known, on the character of the correlations between the fluctuations of the refractive index of the medium. If the fluctuations of the refractive index n of the medium can be described by the correlation function

$$f_n(\vec{r}) = \overline{\Delta n^2} \exp(-r^2/a^2), \tag{5}$$

where $\overline{\Delta n^2} = \overline{(n - \overline{n})^2}$; and a is the inhomogeneity scale, then [1]

$$\rho_E(\vec{l}) = \frac{e^{-l^2/a^2} - \frac{1}{2iD}\left[E_i(\varepsilon_2) - E_i(\varepsilon_1)\right]}{1 - \frac{1}{D}\operatorname{arc\,tg} D}. \tag{6}$$

Here Ei is the integral exponential function; $\varepsilon_1 = \frac{il^2/a^2}{D-i}$;

$\varepsilon_2 = -\frac{il^2/a^2}{D+i}$; $D = 4L/ka^2$; $k = 2\pi/\lambda$; L — length of route

Let us assume that the signal level fluctuations are due exclusively to the transport of the inhomogeneities (i.e., we assume that we can neglect other factors influencing the field fluctuation, such as heat conduction, diffusion, or turbulence). It turns out here that the problem of correlation and time can be reduced to an investigation of the spatial correlation properties of the field amplitude fluctuations [1]. Thus, if the spatial correlation $\rho E(l)$ is described by a formula in the form (6), then the temporal correlation of the field amplitude fluctuations $\rho E(\tau)$ can be represented as

$$\rho_E(\tau) = \frac{e^{-\frac{v_\perp^2\tau^2}{a^2}} - \frac{1}{2iD}\left[E_i(\varepsilon_2) - E_i(\varepsilon_1)\right]}{1 - \frac{1}{D}\operatorname{arctg} D}. \tag{7}$$

The last expression can be obtained from (6) by substituting for l the quantity

$$l_\perp = v_\perp\tau, \tag{8}$$

where $v_\perp$ – is the transverse component of the transport velocity. When D ≫ 1 the expression for $\rho E(\tau)$ simplifies and assumes the form

$$\rho_E(\tau) \simeq e^{-\frac{v_\perp^2\tau^2}{a^2}}. \tag{7'}$$

Introducing the definition of the temporal correlation interval

$$\tau_0 = \frac{a}{v_\perp}, \tag{9}$$

we have

$$\rho_E(\tau) = e^{-\frac{\tau}{\tau_0}}. \tag{7''}$$

It follows from (9) that the correlation interval τ_0 depends on the transverse component of inhomogeneity transport velocity.

An analysis of (6) shows [1] that the correlation when the antennas are transversely separated extends over a distance on the order of the correlation radius of the refractive index of the medium for all three ranges of values of the wave parameter D.

1. EXPERIMENTAL SETUP

The described experimental investigations were carried out over a route 36 km long, extending over a terrain of medium ruggedness. The transmitter pulse power was 65 kw, the pulse repetition rate was 577 pulses/sec, the pulse duration was 1 microsecond, and the carrier frequency was 9,370 Mc. The radiating system was a parabolic antenna 0.7 m in diameter with symmetrical mean) lobe aperture amounting to approximately 3° at the half-power points. The polarization of the radiated waves was horizontal. The height of the transmitting antenna above the earth's surface was approximately 40 meters.

The receiving antennas were pyramidal horns with directivity pattern approximately 15° in the H plane and approximately 18° in the E plane. In many cases rectangular waveguide sections were used as receiving antennas. The antenna was so constructed that the separation between receiving elements could be varied from 4 to 200 cm, and the antennas could be oriented and secured in space. The height of the antennas at the reception point was approximately 4 meters above the earth's surface.

The quantities investigated on the receiving end of the route were the fluctuations of the signal amplitude. Reception was with two superheterodyne receivers with bandwidth 4 Mc and sensitivity 70 dbm. The power amplitude characteristic of the receivers was linear. To check the stability of the receiver system parameters, the characteristic was periodically calibrated against a standard signal generator.

The envelope of the amplitude of the received signal produced at the receiver output was photographed on motion picture film with a MPO-2 oscillograph. Depending on the character of the fluctuations, the film was driven at 2 or 4 mm/sec. Examples of the recorded signals are shown in Figure 1. The resultant oscillograms were processed statistically. The receiver output signal was monitored with a cathode-ray oscilloscope.

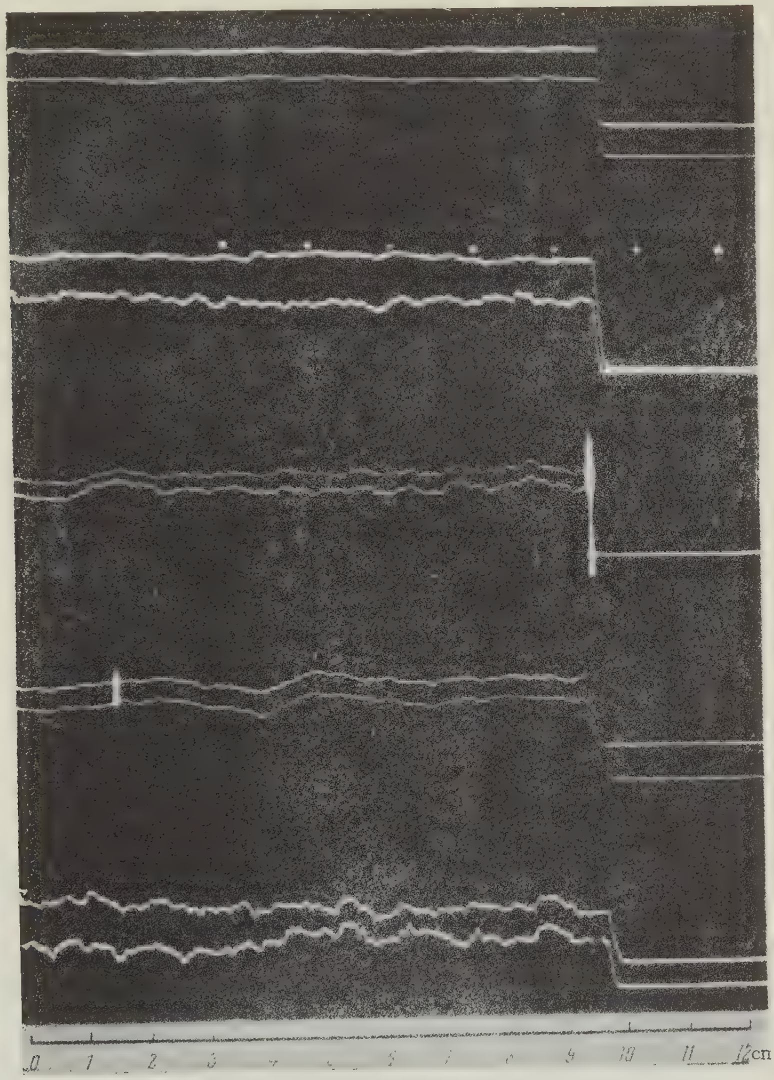


Figure 1. Sample recordings of field level variations.

3. RESULTS OF EXPERIMENTS AND ANALYSIS OF SPATIAL AND TEMPORAL CORRELATION CURVES

The envelopes of the output signals of the two receivers were recorded simultaneously on a common film. The recording time for each separation base was 3 — 4 minutes. To calculate the correlation coefficients, some 180 — 240 pairs of points, spaced $\tau = 1$ sec. apart, were chosen on each portion of film corresponding to a definite separation base. The experimental correlation curves were used to determine for each run the radii of the spatial (l_0) and temporal (τ_0) correlations, taken at the level $\rho_E(l) = \rho_E(\tau) = 0.5$.

a) Analysis of spatial correlation curves. The results of the statistical processing of 20 runs of measurement of the spatial correlation with the antennas horizontally separated transverse to the route are listed in Table 1 and shown in the summary plot, Figure 2.

Table 1

Run	l_0, meters	run	l_0, meters	Run	l_0, meters
29 VI 1959 17^{10}—17^{50}	0.2	16 IX 1959 15^{50}—16^{50}	0.75	11 XI 1959 16^{00}—17^{00}	0.8
30 VI 1959 10^{30}—11^{20}	0.3	17 IX 1959 9^{20}—$1\ ^{20}$	0.5	13 XI 1959 16^{05}—17^{05}	0.2
1 VII 1959 14^{10}—15^{20}	0.4	17 IX 1959 13^{15}—13^{50}	0.2	12 XII 1959 20^{30}—21^{20}	0.8
1 VII 1959 17^{00}—17^{50}	0.2	27 X 1959 15^{30}—16^{30}	0.2	13 XII 1959 5^{45}—6^{35}	0.35
2 VII 1959 10^{00}—10^{50}	0.4	30 X 1959 10^{00}—11^{00}	0.6	13 XII 1959 23^{00}—23^{35}	0.6
2 VII 1959 17^{00}—17^{50}	0.2	30 X 1959 12^{00}—13^{00}	0.5	13—14 XII 1959 23^{40}—0^{30}	0.5
16 IX 1959 12^{00}—13^{00}	0.7	30 X 1959 15^{00}—16^{00}	—	—	—

It is difficult to fit the experimental data of each individual run to some correlation function of definite form for the spread in the experimental points is considerable. A certain idea of the general form of the spatial correlation relationships can be gained from an examination of the summary plot, Figure 2, on which the experimental points, an exponential, and the Gaussian curve are represented by the circle, solid line, and dashed line, respectively. In the construction of both curves, the value of a was chosen to be the radius of the spatial correlation l_0, obtained from experiment. It is seen from Figure 2 that the general course of the distribution of the experimental points is closer to exponentia than to Gaussian. The measurement results, listed in Table 1 and shown in Figure 2, show that the values of the correlation radius at a horizontal spacing of the antenna transverse to the route range from 0.2 to 0.8 meters, i.e., the correlation vanishes as a rule at a distance 70 — 25 wavelengths, confirming the hypothesis advanced in [2] . The average over 20 measurement runs yields for the transverse-correlation radius a value $\overline{l}_0 = 0.5$ m.

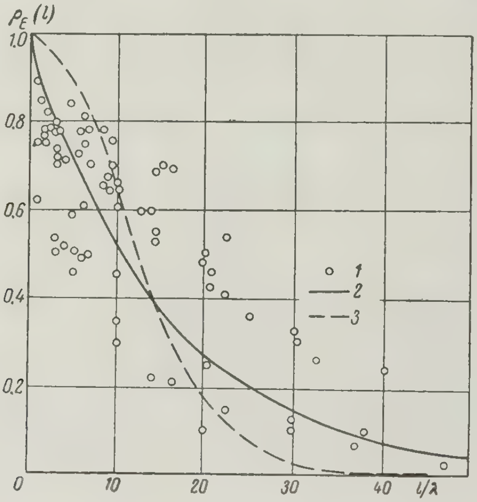


Figure 2. Spatial correlation for transversely spaced receiving antennas: 1 — experimental data; 2 — theoretical curve $\rho_E(l) = \exp(-l/a)$; 3 — theoretical curve $\rho_E(l) = \exp(-l^2/a^2)$.

Inasmuch as very little experimental data on the spatial correlation of the UHF radio signal amplitude fluctuations in direct line of sight visibility have been published to date.

(in addition to the already cited reference [2], we know only of one paper [8], in which experimental values $\rho_3\lambda = 0.97$ and $\rho_{27}\lambda = 0.4$ are given for $\lambda = 7\text{m}$), it is of certain interest to compare these data with the results obtained in UHF tropospheric scatter investigations. Similar information is found, in particular, in references [3, 4, 7], from which it follows that the transverse correlation beyond the horizon vanishes at distances on the order of 10 — 35 wavelengths. A similar comparison enables us, apparently, to assume that the average values of the transverse correlation radius in the propagation of UHF under direct visibility conditions and beyond the horizon, expressed in wavelengths, are quantities of the same order of magnitude.

b) Analysis of the temporal correlation curves. The results of statistical processing and calculations on 26 runs of measurements of the temporal correlation are listed in Table 2 and are illustrated in Figures 3 — 5. An analysis of the course of the temporal correlation curves shows the following.

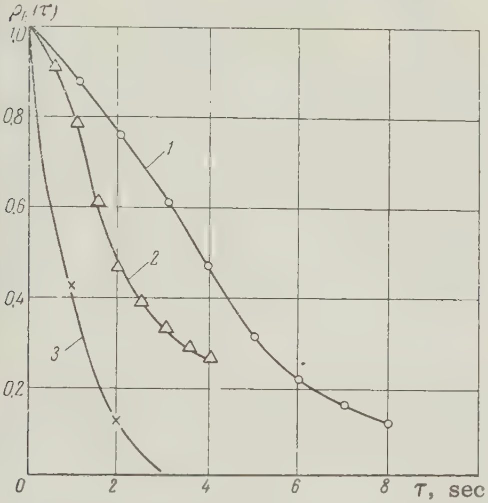


Figure 3. Dependence of rate of decrease in temporal correlation curves on the transverse component of the wind velocity.

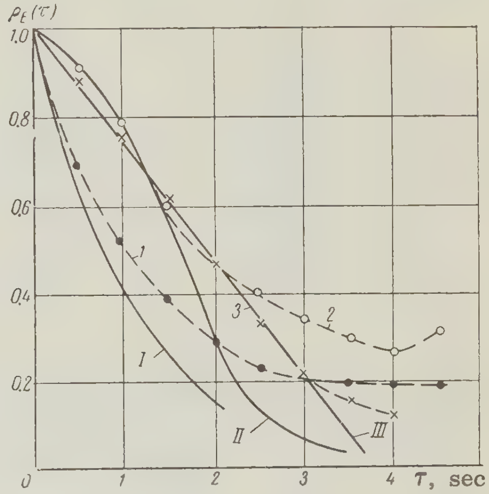


Figure 4. Main types of temporal correlation curves. Experimental curves: 1 — first type (29 June 1959, 17:10 — 17:50); 2 — second type (30 June 1959, 10:30 — 11:20); 3 — third type (3 July 1959, 10:00 — 10:50); theoretical curves:

III — straight line.

1. A clear-cut dependence of the rate of fall-off of the curves on the transverse component of the wind velocity $v_\perp$: the rate of decrease increases with increasing wind velocity. Examples of temporal correlation curves for different values of the transverse component of wind velocity are shown in Figure 3.

2. Three basic types of correlation curves; the course of the first type, within a correlation-coefficient range from 1.0 to 0.5, can be satisfactorily approximated by

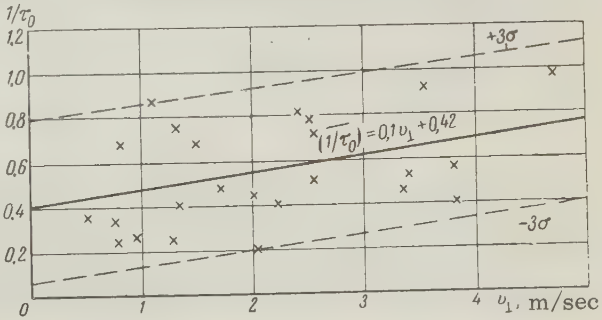


Figure 5. Dependence of transverse component of wind velocity on the reciprocal of the temporal correlation radius.

an exponential curve. The second curve can be approximated with a Gaussian curve and the third with a straight line (within the same limits). Examples illustrating these curves are shown in Figure 4. The theoretical curves on the same figure have been plotted

Table 2

Run	$v_\perp$, m/sec	τ_0, sec	$l_\perp$, m	Run	$v_\perp$, m sec	τ_0, sec	$l_\perp$, m
29 VI 1959 17^{10}—17^{50}	2.5	1.3	3.2	17 IX 1959 9^{20}—10^{20}	3.8	1.75	6.6
30 VI 1959 10^{30}—11^{0}	1.9	1.9	3.6	17 IX 1959 12^{20}—13^{00}	2.9	1.9	5.5
1 VII 1959 14^{10}—15^{20}	4.3	0.8	3.4	17 IX 1959 13^{15}—13^{50}	2.5	2.0	5.0
1 VII 1959 17^{00}—17^{50}	3.5	1.1	3.8	27 X 1959 15^{30}—16^{30}	1.7	2.1	3.5
2 VII 1959 10^{00}—10^{50}	4.7	1.0	4.7	30 X 1959 10^{00}—11^{00}	1.3	1.4	1.8
2 VII 1959 17^{00}—17^{50}	3.0	1.0	3.0	30 X 1959 12^{00}—13^{00}	1.1	1.2	1.3
3 VII 1959 10^{00}—17^{50}	0.9	3.8	3.4	11 XI 1959 16^{00}—17^{00}	0.5	2.9	1.5
3 VII 1959 14^{30}—15^{20}	1.2	4.0	4.8	12 XI 1959 16^{00}—17^{00}	0.8	4.5	3.6
3 VII 1959 17^{00}—17^{50}	0.9	6.0	5.4	13 XI 1959 16^{05}—17^{05}	0.75	3.0	2.2
8 VII 1959 14^{30}—15^{20}	1.3	2.5	3.2	12 XII 1959 20^{30}—21^{20}	0.8	1.5	1.2
10 IX 1959 10^{00}—1^{00}	2.2	2.5	5.5	13 XII 1959 5^{45}—6^{35}	1.1	0.7	0.8
16 IX 1959 12^{00}—13^{00}	2.9	2.2	3.7	13 XII 1959 23^{00}—23^{35}	2.4	1.0	2.4
16 IX 1959 15^{50}—16^{50}	2.0	2.5	5.0	13—14 XII 1959 23^{40}—0^{30}	2.4	0.8	1.9

values of the temporal correlation radius τ_0 taken from the experimental curves of the corresponding runs at the level $\rho_E(\tau) = 0.5$. Most of the experimental curves (17 runs) are close to the first type, four curves are close to the second and five to the third.

As noted earlier, the assumption that the fluctuations of the signal level are due only to the transport of the inhomogeneities enables us to reduce the time-correlation problem to a space-correlation problem [in accordance with formula (8)]. The values of $l_\perp$ calculated for values of τ equal to the radius of the temporal correlation over the level $\rho_E(\tau) = 0.5$. are shown in Table 2 and Figure 5. The latter is a plot of $v_\perp$ vs. $1/\tau_0$. The asterisks denote the experimental points. The line

$$\frac{1}{\tau_0}\Big) = 0.1\, v_\perp + 0.42.$$

which is the regression line of $1/\tau_0$ on $v_\perp$, has been plotted by the method of least squares [5]; the two lines parallel to the regression line delineate a spread in $1/\tau_0$ equal to three times the dispersion ($\pm 3\sigma$).

An analysis of Table 2 and Figure 5 leads to the following:

1. In spite of the considerable scatter in the points, which is mostly the consequence of low accuracy in the measurement of the wind data, a correlation is observed between v and $1/\tau_0$. The presence of such a correlation is indicated also by the dependence of the course of the temporal correlation curves on the transverse wind-velocity component, plotted in Figure 4.

2. The values of $l_\perp$, calculated on the basis of (8) over the level $\rho_E(\tau) = 0.5$, are nearly equal and on the order of several meters. An average over all the 26 runs yields

$$\bar{l}_\perp = 4\,\mathrm{m}.$$

3. Considerable seasonal variations in $l_\perp$ are observed. Thus, whereas the average value of $\bar{l}_\perp$ in June-September is on the order of 4.6 meters, it amounts to two meters in the fall-winter period.

4. The regression lines do not pass through the origin, as expected on the basis of the assumption made concerning the causes of the signal fluctuation. Such a discrepancy

can be qualitatively explained by recognizing that when we determine $l_\perp$ by formula (8) we take into account only the horizontal part of the transverse wind-velocity component and disregard the vertical transfer of air masses due to convection. In addition, considerable contributions to the signal fluctuations may be made also by the longitudinal wind-velocity component [6], and also by the random motion of the inhomogeneities themselves.

c) Comparison of spatial and temporal correlation curves. In a comparison of the values of l_0, obtained from the spatial correlation curves, with the corresponding values l , calculated on the basis of (8) from the temporal curves over the level $\rho_E(\tau) = 0.5$, an appreciable difference between the two is observed, the latter being 8 — 10 times greater than the former. Possible causes of this disparity may be as follows.

1. The effect of the multiple-path field pattern in spatially separated reception.

As noted in [9 — 12], the presence of an interface, which results in an interference field pattern, can cause the experimental radio signal fluctuation characteristics obtained under these conditions to differ greatly from the corresponding values calculated for a medium without boundary.

The character of the route used for the experiment reported here give grounds for assuming that the reflection was negligible over the greatest part of the route, with the exception of a portion of the terrain immediately adjacent to the receiving point.

Making use of published data and taking account of the fact that the transverse correlation radii have been determined from an experimental measurement of the phase at a wavelength $\lambda = 10$ cm [9, 10], it can be assumed that the transverse correlation radii of the amplitude fluctuation ($l \ll \sqrt{\lambda L}$) which we obtain for $\lambda = 3$ cm are in satisfactory agreement with the results of the preceding investigations. We note that no clear-cut interference field pattern was observed in our experiment, for the relative dispersions change by not more than 1.5 when the mobile antenna was displaced within the limits of the maximum separation base.

2. Failure to take into account the random motion of the inhomogeneities themselves; in this case the rate of attenuation of the correlation relations may increase.

In conclusion, I take this pleasant opportunity to express sincere gratitude to A.A. Semenov and G.A. Karpeyev for valuable advice and help in the work, and also to the entire staff of the Department of Electromagnetic Wave Propagation for help in performance of the experiment and processing of the measurement results.

REFERENCES

1. L.A. Chernov, Rasprostraneniye voln v sredye so sluchaynymi neodnorodnostyami, (Propagation of Waves in a Medium with Random Inhomogeneities), Izd. AN SSSR, 1958.
2. A.A. Semenov, G.A. Karpeyev, Vestnik MGU, 1957, 6, 69.
3. H. Staras, Proc. I.R.E., 1955, 43, 10, 1374.
4. J. Kurihara, Proc. I.R.E., 1955, 43, 10, 1363.
5. A.M. Dlin, Matematicheskaya statistika v tekhnikye (Mathematical Statistics in Engineering), Izd. Sovetskaya nauka, 1958.
6. A.D. Wheelon, J. Atmos. and Terr. Phys., 1959, 15, 3/4, 185.
7. J.H. Chisholm, P.A. Portman, J.T. de Bettencourt, J.F. Roche, Proc. I.R.E., 1955, 43, 10, 1317.
8. V.A. Krasil'nikov, Izv. AN SSSR, ser. geograf. i geofiz, 1949, 13, 1, 33
9. A.V. Men', V.I. Gorbach, S.Ya. Braude, Izv. vuzov MVO SSSR, Radiofizika, 1959, 2, 3, 388.
10. A.V. Men', S.Ya. Braude, V.I. Gorbach, Dokl. AN SSSR, 1959, 125, 5, 1019.
11. E.A. Kanner, F.G. Bass, Izv. vuzov MVO SSSR (Radiofizika), 1959, 2, 4, 553.
12. F.G. Bass, E.A. Kanner, Izv. vuzov MVO SSSR (Radiofizika), 1959, 2, 4, 565.

Physics Faculty, Moscow State University Submitted to the editors 6 May, 1960.

STATISTICAL ANALYSIS OF OPTIMAL RESOLUTION

Ya.D. Shirman

The possibility is investigated of resolving between primary and secondary radio-wave emitters in the presence of noise from thermal radio emission. It is assumed that the resolution is based on data concerning the temporal field distribution on a segment of a straigh line and that the dimension of the segment is large compared with the maximum wavelength of the investigated oscillations. The coefficient of energy utilization and the law of optimal processing in resolution are found. Examples are given of the application of the analysis results to the case of angular resolution.

INTRODUCTION

The operation of many electronic devices can be described as the detection of primary or secondary radio wave emitters. We shall assume that two objects to be detected are resolved if the presence of one of these objects does not greatly deteriorate the possibility of detecting the second object. One of the causes of difficulty in determining accurately the distribution of radiating objects is the presence of thermal radiation, which limits the possibilities of detection and resolution of objects even if the receiver is practically noiseless. In view of the random character of the interfering oscillations and of the individual parameters of the resolved signals (for example, amplitude and phase), the problems of detection and resolution are statistical, and the operation of devices for optimal detection and resolution reduces to an analysis of the distribution of the field components at different points of the antenna aperture as functions of the time.

Thus, not only the receiver but the antenna itself are considered as an element of a computer that solves the problem of the most probable distribution of radiators. This point of view is illustrated below by means of a very simple example. It is assumed that the resolution is effected by an analysis of the field components as functions of the time within a stationary segment of the line $-l/2 < x < l/2$. Each of the field components is the result of a superposition of a δ-correlated thermal radiation and radiation from the objects to be resolved. (the effect of the internal receiver noise can be accounted for as a corresponding increase in the thermal-radiation temperature). The dimension l is considered to be sufficiently large compared with the maximum wavelength of the radiated oscillations and small compared with the distance to the radiators.

1. DISCRETIZATION OF THE OSCILLATIONS ON THE APERTURE

Inasmuch as each of the analyzed field components is not only a function of the time t, as in the well-known theory of optimal reception, but also of the coordinates x of the points of the aperture, it is expedient to carry out discretization over the coordinate x or its spectru along with discretization in time or over the frequency spectrum. We give below the discretization over the frequency f of the function of the time and over the coordinate x of the points of the aperture.

Assuming that any of the analyzed high-frequency functions u(t, x) is specified only for a limited time interval T, we describe this function on this interval by a sum of Fourier-series harmonics

$$u(t, x) = \sum_m u_m(t, x) = \sum_m [a_m(x) \cos 2\pi f_m t + b_m(x) \sin 2\pi f_m t], \tag{1}$$

corresponding to the discrete frequencies

$$f_m = \frac{m}{T} \; (m = m_1, m_1 + 1, \ldots, m_2). \tag{2}$$

If the distribution u(t, x) is produced by a system of s discrete radiators, spaced sufficiently far apart to make the waves from them plane within the limits of the aperture, then for each of the field harmonics $u_m(t, x)$, we can write

$$u_m(t,x)=\sum_{i=1}^{s}u_{mi}(t,x)=\sum_{i=1}^{s}u_{mi}\left(t-\frac{x}{c}\cos\alpha_i,\,0\right)=$$
$$=\sum_{i=1}^{s}\left[a_{mi}(0)\cos 2\pi\left(f_m t-\frac{f_m\cos\alpha_i}{c}x\right)+b_{mi}(0)\sin 2\pi\left(f_m t-\frac{f_m\cos\alpha_i}{c}x\right)\right]. \tag{3}$$

Here the index i characterizes the i-th radiator; c is the velocity of light and α_i is the angle between the direction of arrival of the wave from the i-th radiator and the positive direction of the x-axis.

Comparing (1) and (3) we can get

$$a_m(x)=\sum_{i=1}^{s}\left[a_{mi}(0)\cos 2\pi\frac{f_m\cos\alpha_i}{c}x-b_{mi}(0)\sin 2\pi\frac{f_m\cos\alpha_i}{c}x\right],$$
$$b_m(x)=\sum_{i=1}^{s}\left[a_{mi}(0)\sin 2\pi\frac{f_m\cos\alpha_i}{c}x+b_{mi}(0)\sin 2\pi\frac{f_m\cos\alpha_i}{c}x\right]. \tag{4}$$

Expressions (4) can be regarded as spectral expansions of the functions $a_m(x)$ and $b_m(x)$ of the coordinate x. As follows from the foregoing expressions, the frequency spectrum over the coordinate x is discrete and contains the frequencies $f_m|\cos\alpha_i|/c$. Not one of the frequencies of the spectrum exceeds f_m/c. This last conclusion remains in force if we change over to a continuous distribution of part of the radiators over the angles α, as takes place in thermal radiation, and thereby include a continuous spectrum component along with the discrete one.

It follows from the foregoing analysis that the functions $a_m(x)$ and $b_m(x)$, which enter into (1), have a limited frequency spectrum over the coordinate x, ranging from 0 to f_m/c. One can apply to such functions the Kotel'nikov theorem, i.e.,

$$a_m(x)=\sum_n a_m(x_{mn})\frac{\sin 2\pi\frac{f_m}{c}(x-x_{mn})}{2\pi\frac{f_m}{c}(x-x_{mn})},$$
$$b_m(x)=\sum_n b_m(x_{mn})\frac{\sin 2\pi\frac{f_m}{c}(x-x_{mn})}{2\pi\frac{f_m}{c}(x-x_{mn})}. \tag{5}$$

Here x_{mn} are the coordinates of the points within the limits of the entire x-axis

$$x_{mn}=\frac{n}{2}\frac{c}{f_m}. \tag{6}$$

Relations (5) indicate the possibility of spatial discretization of the coefficients of an arbitrary n-th harmonic. The discretization points x_{mn} are arranged with a spacing $c/2f_m = \lambda_m/2$ where λ_m is the wavelength corresponding to the m-th harmonic. If the discretization interval is small compared with the dimension of the aperture l, then we can describe the field within its confines by neglecting the edge effects and limiting ourselves only to the first terms of the series, for which the discretization points are located within the confines of the aperture.

Combining (1) and (5) we get

$$u(t,x)=\sum_{m,n}\left[a_{mn}\cos 2\pi f_m t+b_{mn}\sin 2\pi f_m t\right]\frac{\sin 2\pi\frac{f_m}{c}(x-x_{mn})}{2\pi\frac{f_m}{c}(x-x_{mn})}, \tag{7}$$

where

$$a_{mn}=a_m(x_{mn});\quad b_{mn}=b_m(x_{mn}). \tag{8}$$

2. LAW OF DISTRIBUTION OF THE REALIZATIONS OF THE THERMAL RADIATION FIELD COMPONENTS IN TIME AND OVER THE ANTENNA APERTURE

The probabilities of different realization u(t, x) of the thermal-radiation field components can be described by a law of joint distribution of the coefficients a_{mn} and b_{mn}, each of which is a Gaussian random quantity with zero mean value. This law is fully characterized by specifying the dispersions $\overline{a^2_{mn}}, \overline{b^2_{mn}}$ and the mixed second moments of distribution (covariations) $\overline{a_{mn}\, a_{\mu\nu}}, \overline{b_{mn}\, b_{\mu\nu}}, \overline{a_{mn} b_{\mu\nu}}$ (the superior bar denotes averaging, and m, n, μ, and ν are integers).

Assuming the spectral density of the thermal radiation to be constant within the working frequency band, $N_0(f) = N_0$, and taking into account the interval $\Delta f_m = 1/T$ between the spectral harmonics, we put $0.5\ (\overline{a^2_{mn}} + \overline{b^2_{mn}}) = N_0(f_m)\ \Delta f_m$, hence

$$\overline{a^2_{mn}} = \overline{b^2_{mn}} = \overline{a^2_m(x_{mn})} = \overline{b^2_m(x_{mn})} = \frac{N_0}{T}. \tag{9}$$

Inasmuch as the thermal radiation can be assumed to be uncorrelated over the spectrum, the covariations with different harmonic numbers $\mu \neq m$ vanish identically. To estimate the covariations with numbers $\mu = m$ and with discretization indices n and ν, we make use of formulas (4). First subdividing the entire sphere into small solid angles $\Delta\Omega_i = \sin\alpha_i\Delta\alpha_i\Delta\beta_i$ α and β are spherical coordinates) and assuming the thermal radiation to be uniformly distributed over the solid angles within the limits of the entire space, we introduce equivalent discrete radiators with field parameters at the aperture center

$$\overline{a^2_{mi}(0)} = \overline{b^2_{mi}(0)} = \frac{N_0}{T}\frac{\Delta\Omega_i}{4\pi} = \frac{N_0}{4\pi T}\sin\alpha_i\Delta\alpha_i\Delta\beta_i. \tag{10}$$

By virtue of the independent character of the radiators with different indices i and j, and also the independence of the quadrature components of one and the same radiator, we have

$$\overline{a_{mi}(0)\, a_{mj}(0)} = \overline{b_{mi}(0)\, b_{mj}} = \overline{a_{mi}(0)\, b_{mj}(0)} = 0 \quad \text{for} \quad i \neq j. \tag{11}$$

Calculating the products of expressions similar to (4) and averaging with allowance for relations (10) and (11), we obtain after very simple transformations

$$\overline{a_{mn}a_{m\nu}} = \overline{b_{mn}b_{m\nu}} = \frac{N_0}{4\pi T}\sum_{i=1}^{s}\cos\left[2\pi\frac{f_m\cos\alpha_i}{c}(x_{mn} - x_{m\nu})\right]\sin\alpha_i\Delta\alpha_i\Delta\beta_i.$$

Passing to the limit of the corresponding double integral and calculating its value, we get

$$\frac{1}{4\pi}\int_0^{2\pi} d\beta\int_0^{\pi}\cos\left[2\pi\frac{f_m\cos\alpha}{c}(x_{mn} - x_{m\nu})\right]\sin\alpha d\alpha = \frac{\sin\left[2\pi\frac{f_m}{c}(x_{mn} - x_{m\nu})\right]}{2\pi\frac{f_m}{c}(x_{mn} - x_{m\nu})}.$$

Using (6), we obtain finally

$$\overline{a_{mn}a_{m\nu}} = \overline{b_{mn}b_{m\nu}} = \frac{N_0}{T}\frac{\sin\pi(n-\nu)}{\pi(n-\nu)}, \tag{12}$$

from which it follows that $\overline{a_{mn}a_{m\nu}}$ and $\overline{b_{mn}b_{m\nu}}$ vanish when $n \neq \nu$. We can analogously verify that $\overline{a_{mn}b_{m\nu}} = 0$ both when $n \neq \nu$ and when $n = \nu$.

We thus establish that the expansion coefficients a_{mn} and b_{mn} of the thermal-radiation field components are independent random quantities with dispersion M_0/T and A zero mean value.

The joint probability density of the totality of these quantities can be represented in the form

$$p(\ldots, a_{mn}, b_{mn}, \ldots) = \left(\frac{T}{2\pi N_0}\right)^{0{,}5\nu_0} e^{-\frac{1}{N_0}\left[\left(a_{mn}\sqrt{\frac{T}{2}}\right)^2 + \left(b_{mn}\sqrt{\frac{T}{2}}\right)^2\right]}, \tag{13}$$

where ν_0 is the total number of the independent coefficients of the series.

We shall agree to regard the parameters

$$a_{mn}\sqrt{\frac{T}{2}} \text{ and } b_{mn}\sqrt{\frac{T}{2}} \tag{14}$$

of the arbitrary function u(t, x) as coordinates $u_k (k = 1, 2, \ldots)$ of a certain multidimensional vector u. Defining the scalar product of the two multi-dimensional vectors as

$$\vec{u}\,\vec{v} = \sum_k u_k v_k \tag{15}$$

and putting $\vec{u}\vec{u} = \vec{u}^2$, we can write (13) in shorter form after making a change of variables

$$p\,(\vec{u}) = (\pi N_0)^{-0{,}5\nu_0} e^{-\frac{\vec{u}^2}{N_0}}. \tag{16}$$

In order to reduce (15) to an integral form, we introduce the corresponding spectral densities, for example

$$g_u\,(f,\,x) = \int\limits_{-\infty}^{\infty} u\,(t,\,x)\,e^{-j2\pi ft}dt, \tag{17}$$

where the function u(t, x) differs from zero only for the time interval T. As indicated in the appendix, if the aperture dimension l is sufficiently large compared with the maximum wavelength $\lambda_{\max}$ and for an observation interval $T \to \infty$, the following approximate relation holds true:

$$\vec{u}\,\vec{v} = \frac{4}{c}\,\mathrm{Re}\int\limits_0^{\infty} df \int\limits_{-l/2}^{l/2} f g_u\,(f,\,x)\,g_v^*\,(f,\,x)\,dx. \tag{18}$$

It is all the more accurate, the greater the aperture l compared with $\lambda_{\max}$ and, consequently, the less influential the edge effect is in the use of the Kotel'nikov theorem.

If the signal spectrum is concentrated around the frequency f_0, the nonoscillating factor $f \simeq f_0 = c/\lambda_0$ can be taken outside the integral sign. Interchanging the order and the limits of integration, we obtain

$$\vec{u}\,\vec{v} = \frac{2}{\lambda_0}\int\limits_{-l/2}^{l/2} dx \int\limits_{-\infty}^{\infty} g_u\,(f,\,x)\,g_v^*\,(f,\,x)\,df. \tag{19}$$

Using the Parseval theorem and going to complex amplitudes, we obtain

$$\vec{u}\,\vec{v} = \frac{2}{\lambda_0}\int\limits_{-l/2}^{l/2} dx \int\limits_{-\infty}^{\infty} u\,(t,\,x)\,v\,(t,\,x)\,dt = \frac{1}{\lambda_0}\,\mathrm{Re}\int\limits_{-l/2}^{l/2} dx \int\limits_{-\infty}^{\infty} \dot{U}\,(t,\,x)\,\dot{V}^*\,(t,\,x)\,dx. \tag{20}$$

3. STATISTICAL ANALYSIS OF OPTIMAL RESOLUTION

The resolution can be reduced to observation of the useful signal against the background of thermal noise and interfering signals with random amplitudes and random initial phases. Signals from neighboring objects, which make detection difficult, can be regarded as interfering. The corresponding problem has been solved in [1] as applied to functions of time (not of the coordinate and time) for a Rayleigh amplitude distribution and equiprobably distribution of initial phases of the interfering signals. Inasmuch as the initial relation (15) of the present paper coincides with relation (16) of reference [1], and the problem can be identically formulated in vector form for both cases, all the vector relations of reference [1] are fully applicable to the problem solved in the present article. The integral form of the relations, however, is changed, since the expressions for the scalar products of multidimensional vectors have in this case a somewhat more complicated notation (double integrals instead of single ones).

Thus, for example, in [1] we introduced a useful-signal-energy utilization coefficient k, for detection against the background of interfering signals and noise. This coefficient is equal to the ratio of the energy of the threshold signal in this case to the corresponding energy in detection against the background of noise only, for fixed values of the probability of correct detection, D, and the probability of false alarm, F. The latter is understood to be "pointlike" and not "interval-like," i.e., for a fixed time of signal arrival, for a fixed angular direction, etc. Taking into account all the foregoing considerations, we obtain the expression for k in the form

$$k = \frac{\vec{u}_{m+1}\vec{r}_{m+1}}{\vec{u}^2_{m+1}} = \frac{\iint \dot{U}_{m+1}(t, x)\, \dot{R}^*_{m+1}(t, x)\, dt\, dx}{\iint |\dot{U}_{m+1}(t, x)|^2\, dt\, dx}. \tag{21}$$

Here the complex amplitude $\dot{R}_{m+1}(t,x)$ is determined from the following recurrence relation:

$$\dot{R}_{i+1}(t, x) = \dot{U}_{i+1}(t, x) - \sum_{j=1}^{i} \dot{R}_j(t, x) \frac{\iint \dot{U}_{i+1}(t, x)\, \dot{R}^*_j(t, x)\, dt\, dx}{\frac{\lambda_0}{2} N_0 + \iint \dot{U}_j(t, x)\, \dot{R}^*_j(t, x)\, dt\, dx} \qquad (i = 1, \ldots, m). \tag{22}$$

We bear in mind here that $\dot{R}_1(t,x) = \dot{U}_1(t,x)$. The function $\dot{R}_{m+1}(t,x)$ satisfies also the integral equation

$$\dot{R}_{m+1}(t, x) + \iint \varphi(s, y, t, x)\, R_{m+1}(s, y)\, ds\, dy = \dot{U}_{m+1}(t, x) \tag{23}$$

with degenerate kernel

$$\varphi(s, y, t, x) = \frac{2}{N_0 \lambda_0} \sum_{j=1}^{m} \dot{U}_j(t, x)\, \dot{U}^*_j(s, y), \tag{24}$$

which can be of interest in a consideration of the case $m \to \infty$.

For a signal with random initial phase and nonrandom ampoitude (or an amplitude which is random and has a Rayleigh distribution) the optimal processing of the received oscillation $\dot{U}(t,x)$ reduces to a calculation of the quantity

$$\left| \iint \dot{U}(t, x)\, R^*_{m+1}(t, x)\, dt\, dx \right| \tag{25}$$

and subsequent comparison with some threshold.

If only two signals are resolved and the noise level is low compared with the interfering signal

$$N_0 \ll 2\vec{u}^2_{\text{int}} = \frac{2}{\lambda_0} \iint |U_{\text{int}}(t, x)|^2\, dt\, dx$$

($\vec{u}_{\text{int}}$ is respectively either $\vec{u}_1$ or $\vec{u}_2$) the energy utilization coefficient will be

$$k \simeq 1 - \rho. \tag{26}$$

Here ρ is the coefficient of "correlation" of the signals to be resolved:

$$\rho = \frac{\left| \iint \dot{U}_1(t, x)\, \dot{U}^*_2(t, x)\, dt\, dx \right|^2}{\iint |\dot{U}_1(t, x)|^2\, dt\, dx \iint |\dot{U}_2(t, x)|^2\, dt\, dx}, \tag{27}$$

which are described in this case by the nonrandom functions $\dot{U}_1(t,x)$ and $\dot{U}_2(t,x)$.

If two or three components of three-dimensional vectors are used for detection, the products of the scalar quantities in relations (21) — (27) should be replaced by scalar products of the field vectors.

4. EXAMPLES OF APPLICATION OF THE RESULTS OF THE ANALYSIS

Assume that we are resolving two radar signals with complex amplitudes

$$\dot{U}_1(t, x) = \dot{A}\left(t - \tau_1 - \frac{x}{c}\cos\alpha_1\right) e^{-j\Omega_1 t} e^{-j\frac{2\pi}{\lambda_0} x \cos\alpha_1}, \tag{28}$$

$$\dot{U}_2(t, x) = \dot{A}\left(t - \tau_2 - \frac{x}{c}\cos\alpha_2\right) e^{-j\Omega_2 t} e^{-j\frac{2\pi}{\lambda_0} x \cos\alpha_2}, \tag{29}$$

corresponding to nonmonochromatic plane waves, which differ in angle of arrival α, in the in the delay time τ, and in the Doppler frequency Ω. Here we have by virtue of (27) — (29)

$$\rho = \rho(\tau_1, \tau_2, \Omega_1, \Omega_2, \alpha_1, \alpha_2). \tag{30}$$

If the delays on the aperture are small compared with the reciprocal of the frequency-spectrum width, i.e.,

$$\dot A\left(t-\tau_1-\frac{x}{c}\cos\alpha_1\right)\simeq\dot A\,(t-\tau_1),\quad \dot A\left(t-\tau_2-\frac{x}{c}\cos\alpha_2\right)\simeq\dot A\,(t-\tau_2),$$

then

$$\rho=\rho_{\mathrm{I}}\,(\tau_2-\tau_1,\ \Omega_2-\Omega_1)\,\rho_{\mathrm{II}}\,(\alpha_1,\alpha_2), \tag{31}$$

where $\rho(\Delta\tau,\Delta\Omega)$ is the range and velocity auto-correlation function, introduced by Woodward [2] and Siebert [3] (independent of the energy utilization coefficient k) ; $\rho_{\mathrm{II}}(\alpha_1,\alpha_2)$ is a new autocorrelation function of the angle coordinates. Here

$$\rho_{\mathrm{I}}\,(\Delta\tau,\ \Delta\Omega)=\frac{\left|\int\limits_{-\infty}^{\infty}\dot A\,(t)\,\dot A^*\,(t-\Delta\tau)\,e^{j\Delta\Omega t}\,dt\right|^2}{\left[\int\limits_{-\infty}^{\infty}|\dot A\,(t)|^2\,dt\right]^2} \tag{32}$$

and

$$\rho_{\mathrm{II}}\,(\alpha_1,\alpha_2)=\frac{\sin^2\frac{\pi l}{\lambda_0}(\cos\alpha_2-\cos\alpha_1)}{\left[\frac{\pi l}{\lambda_0}(\cos\alpha_2-\cos\alpha_1)\right]^2}. \tag{33}$$

Let us estimate the coefficient of energy utilization when the resolution is only by the angle coordinates ($\tau_2=\tau_1$, $\Omega_2=\Omega_1$, $\rho_{\mathrm{I}}=1$, $\rho=\rho_{\mathrm{II}}$). Putting $\Delta\alpha=\alpha_2-\alpha_1\ll\pi$, we first rewrite (33) in the form

$$\rho_{\mathrm{II}}=\rho_{\mathrm{II}}\,(\Delta\alpha)=\frac{\sin^2(\pi\Delta\alpha/\alpha_0)}{(\pi\Delta\alpha/\alpha_0)^2}, \tag{34}$$

where $\alpha_0=\lambda_0/l\,\sin\alpha$ ($\alpha=(\alpha_1+\alpha_2)/2$). The quantity $2\alpha_0$ can be called the width of the equivalent directivity characteristic of the aperture at the zeros (without allowance for the interfering radiators). Owing to the increase in the dimension l, and consequently the reduction in the width of the equivalent characteristic $2\alpha_0$, we can reduce ρ_{II} to values much less than unity. This corresponds to making it possible to resolve objects with practically no loss in the energy utilization coefficient, something extensively used in practice.

In individual cases it is possible to obtain angular resolution without increasing the dimension l, namely to obtain resolution within the width of the equivalent characteristic through reducing the energy utilization coefficient $k\simeq1-\rho=1-\rho_{\mathrm{II}}$. Retaining only the first two terms of expression (34) for ρ_{II} in the expansion in powers of $\pi\Delta\alpha/\alpha_0$, we obtain

$$k\simeq\frac{1}{3}\left(\frac{\pi\Delta\alpha}{\alpha_0}\right)^2. \tag{35}$$

This means that the presence of reflector 1 does not interfere with a correct deduction of the presence or absence of reflector 2 (the mutual placement of the reflectors is known) provided a signal-energy margin of a factor of 1/k is available. When $\Delta\alpha/\alpha_0$ we have $k\simeq0.13$ and $1/k\simeq7.5$.

Let us analyze the character of the corresponding optimal processing, by which one can reduce the action of an interfering signal which is sufficiently strong compared with the thermal-noise level ($N_0\to0$). From (22), (28), and (29) we obtain for this case

$$\dot R_2\,(t,x)=\dot A\,(t)\left\{e^{-j\frac{2\pi}{\lambda_0}x\cos\alpha_2}-\frac{\sin\left[\frac{\pi l}{\lambda_0}(\cos\alpha_2-\cos\alpha_1)\right]}{\frac{\pi l}{\lambda_0}(\cos\alpha_2-\cos\alpha_1)}\,e^{-j\frac{2\pi}{\lambda_0}x\cos\alpha_1}\right\}. \tag{36}$$

The function $\dot R_2(t,x)$ corresponds to a completely defined processing of the elementary plane waves

$$\dot e^{\,j2\pi f\left(t-\frac{x}{c}\cos\alpha\right)}=\dot U_{f,\,\alpha}\,(t,x)\,e^{j2\pi f_0 t},$$

which in accordance with rule (25) is characterized by the relation

$$K_2(f, \alpha) = \iint \dot{R}_2^*(t, x)\, e^{j2\pi\left[(f-f_0)t - f\frac{x}{c}\cos\alpha\right]}\, dt\, dx. \tag{37}$$

The quantity $K_2(f, \alpha)$, determined by expressions (36) and (37), is a product of the optimal frequency characteristic K(f) and the optimal directivity characteristic $\psi(\alpha)$. For $f \simeq f_0$ we have

$$\psi(\alpha) = \frac{\sin\left[\frac{\pi l}{\lambda_0}(\cos\alpha - \cos\alpha_2)\right]}{\frac{\pi l}{\lambda_0}(\cos\alpha - \cos\alpha_2)} - \frac{\sin\left[\frac{\pi l}{\lambda_0}(\cos\alpha - \cos\alpha_1)\right]}{\frac{\pi l}{\lambda_0}(\cos\alpha - \cos\alpha_1)} \frac{\sin\left[\frac{\pi l}{\lambda_0}(\cos\alpha_2 - \cos\alpha_1)\right]}{\frac{\pi l}{\lambda_0}(\cos\alpha_2 - \cos\alpha_1)}.$$

This characteristic ensures zero reception from the direction of the interfering object, and this indeed provides the resolution.

Let us proceed to a case when the placement of the interfering reflectors is not known exactly, and it is necessary to assume that they are uniformly distributed over the angle coordinates. Taking into account only the differences in the angle coordinates, we put

$$\dot{U}_{m+1}(t, x) = \dot{U}_0(t, x) = \dot{A}_0(t)\, e^{-j\frac{2\pi}{\lambda_0}x\cos\alpha_0}$$

and

$$\dot{U}_j(t, x) = \dot{A}_0(t)\, e^{-j\frac{2\pi}{\lambda_0}x\cos\alpha_j} \sqrt{\frac{d}{4\pi}\sin\alpha_j \Delta\alpha_j \Delta\beta_j},$$

where d characterizes the ratio of the volume energy densities of the summary interference of the useful signal, while $j \leq m$.

Changing from the sum (24) to an integral, and integrating over α and β, we obtain

$$\varphi(s, y, t, x) = \frac{2d}{N_0}\dot{A}_0(t)\,\dot{A}_0^*(s)\,\frac{\sin\frac{2\pi}{\lambda_0}(x-y)}{\frac{2\pi}{\lambda_0}(x-y)}.$$

Inasmuch as $\varphi(s, y, t, x)$ and $\dot{U}_{m+1}(t, x)$ are proportional to $A_0(t)$, we can represent the solution of (23) in the form $\dot{R}(t, x) = A_0(t)\, r(x)$, where r(x) satisfies the integral equatio

$$r(x) + \frac{2\varkappa}{\lambda_0}\int_{-l/2}^{l/2} \frac{\sin\frac{2\pi}{\lambda_0}(x-y)}{\frac{2\pi}{\lambda_0}(x-y)}\, r(y)\, dy = e^{-j\frac{2\pi}{\lambda_0}x\cos\alpha_0},$$

here

$$\varkappa = \frac{d}{N_0}\int_{-\infty}^{\infty} |\dot{A}_0(s)|^2\, ds.$$

For the case of a rather long antenna $\lambda_0/l \to 0$, the solution of the integral equation can be found in the form

$$r(x) = \frac{1}{1+\varkappa}\, e^{-j\frac{2\pi}{\lambda_0}x\cos\alpha_0},$$

which corresponds to an optimal directivity characteristic

$$\psi(\alpha) = \frac{\sin\left[\frac{\pi l}{\lambda_0}(\cos\alpha - \cos\alpha_0)\right]}{\frac{\pi l}{\lambda_0}(\cos\alpha - \cos\alpha_0)},$$

the same as in the absence of interfering reflectors.

This means that any artificial measures (without increasing the aperture) cannot impro the resolution in this case.

The approach to resolution problems, developed in the present article, is not the only one possible. The detected object can be regarded as resolved if the presence of another object (or a series of other objects) does not appreciably deteriorate the accuracy of the parameter measurement. The main formulas of this paper can be extended to include the case when, for example, we use in lieu of the likelihood-ratio analysis the analysis of the a posteriori probability of the values of the measured parameters. The extension of these results to include apertures of antennas with discontinuities entails no great difficulty provided the lengths of the gaps is small compared with the wavelength λ_0; this extension reduces to substitution of suitable integration limits in the formulas of Section 3. One can obviously also lift the assumption of the sufficiently large dimension of the aperture (compared with λ_0) and its one-dimensional nature, but for this purpose it is necessary to changeover to a more general form of canonical expansions than given by the Kotel'nikov theorem.

APPENDIX

Replacing in (17) the function u(t, x) by the series (7) within the limits of the time interval T, and by zero outside these limits, we obtain for $f = f_m$

$$g_u(f_m, x) = \frac{T}{2}\sum_n (a_{mn} - jb_{mn}) \frac{\sin\left[\frac{2\pi}{c} f_m (x - x_{mn})\right]}{\frac{2\pi}{c} f_m (x - x_{mn})} =$$

$$= \sum_n g_u(f_m, x) \frac{\sin\left[\frac{2\pi}{c} f_m (x - x_{mn})\right]}{\frac{2\pi}{c} f_m (x - x_{mn})}. \tag{38}$$

Inasmuch as we have for the function u(t, x)

$$a_{mn} = 2\,\mathrm{Re}\, g_u(f_m, x_{mn})/T, \quad b_{mn} = -2\,\mathrm{Im}\, g_u(f_m, x_{mn})/T \tag{39}$$

[and analogously for the function v(t, x)], using the equality Re (AB) = Re A Re B + + Im A Im B and the relations (14), (15), and (39) we get

$$\overrightarrow{u}\overrightarrow{v} = \frac{2}{T}\,\mathrm{Re}\sum_{m,\,n} g_u(f_m, x_{mn})\, g_v^*(f_m, x_{mn}). \tag{40}$$

By virtue of (38), the following equality holds

$$\int_{-l/2}^{l/2} g_u(f_m, x)\, g_v^*(f_m, x)\, dx =$$

$$= \sum_{n,\nu} g_u(f_m, x_{mn})\, g_v^*(f_m, x_{mn}) \int_{-l/2}^{l/2} \frac{\sin\left[\frac{2\pi}{c} f_m (x - x_{mn})\right]}{\frac{2\pi}{c} f_m (x - x_{mn})} \frac{\sin\left[\frac{2\pi}{c} f_m (x - x_{m\nu})\right]}{\frac{2\pi}{c} f_m (x - x_{m\nu})}\, dx, \tag{41}$$

and for infinite integration limits ($l \to \infty$)

$$\int g_u(f_m, x)\, g_v^*(f_m, x)\, dx = \frac{c}{2f_m}\sum_n g_u(f_m, x_{mn})\, g_v^*(f_m, x_{mn}). \tag{42}$$

We note that the integrals in the right half of (41) are determined by the values of the integrand on the segment between the closest discretization points which include the points x_{mn} and $x_{m\nu}$. We can approximately use relation (42) also for finite l, if the summation is carried out over discretization point located within the limits of the aperture, i.e., the same as in (7). (38), and (40).

Substituting $1/T = \Delta f_m \to 0$, we obtain $\overrightarrow{u}\overrightarrow{v} = \frac{4}{c}\,\mathrm{Re}\int_0^\infty df \int_{-l/2}^{l/2} f g_u(f, x)\, g_v^*(f, x)\, dx.$

REFERENCES

1. Ya.D. Shirman, Theory of Detection of Useful Signal Against the Background of Gaussian Noise and An Arbitrary Number of Interfering Signals With Random Amplitudes and Random Initial Phases, Radiotekhnika i elektronika, 1959, 4, 12, 1
2. P.M. Woodward, Teoriya veroyatnostey i informatsii s primeneniyami v radiolokatsii, (trans. from English) Izd. Sovyetskoye radio, 1955.
3. Siebert, General Laws of Target Detection with the Aid of Radar, (Translation) Problemy radiolokatsionnoi tekhniki (Problems in Radar Techniques), 1957, 5, (41), 3.

Submitted to the editors 24 September 1960

TRANSVERSE DIFFUSION IN DIFFRACTION BY AN IMPEDANCE CYLINDER OF LARGE RADIUS. PART I. PARABOLIC EQUATION IN RAY COORDINATES

G.D. Malyuzhinets and L.A. Vaynshteyn

The two-dimensional problem is considered of diffraction of a cylindrical or plane wave on a circular cylinder, with boundary conditions of the impedance type (under the condition that the cylinder radius is large compared with the wavelength). This problem is solved with the aid of the Leontovich-Fok parabolic equation, written in ray coordinates. This equation takes into account the "transverse diffusion" of the wave amplitude. A general solution is obtained for the diffraction problem, suitable for all distances from the source an from the point of observation to the cylinder and agreeing with the geometrical-optics formulas in the illuminated region.

INTRODUCTION

M.A. Leontovich and V.A. Fok [1 — 4] proposed in their basic papers a parabolic-equation method for the investigation of the asymptotic laws of diffraction; this method was used to solve the problem of propagation of radio waves along the flat [1, 2] and spherical [2] earth's surface, and also the problem of the field of a plane electromagnetic wave in the penumbral region near the surface of a conducting convex body [3, 4], the dimensions and radii of curvature of which are considerably greater than the wavelength. It is found, however, that the parabolic equation, written in beam (involute) coordinates and accounting for only the transverse diffusion of the wave amplitude, permits a determination of the diffraction field in the entire shadow region at arbitrary distances from the convex body, and also in the penumbral region — up to convergence to the formulas of geometric optics. In this paper we demonstrate this fact, using as an example the diffraction of cylindrical and plane waves on an impedance circular cylinder, the radius of which is much greater than the wavelength.

The theoretical treatment of diffraction problems, used in the present paper, was reported several times by G.D. Malyuzhinets in papers delivered to conferences in 1946, 1957, 1959, and 1960 (see [5 — 8]), and also in a review article [9]. This treatment is quite general and, as will be shown later, includes the concept of diffraction rays.

1. MANY-SHEETED PLANE AND RAY COORDINATES

The two-dimensional problem of the diffraction of a cylindrical wave on a circular cylinder of radius r reduces to finding the Green's function G, satisfying the wave equation

$$\frac{\partial^2 G}{\partial r^2} + \frac{1}{r}\frac{\partial G}{\partial r} + \frac{1}{r^2}\frac{\partial^2 G}{\partial \varphi^2} + k^2 G = 0 \quad (a < r < \infty) \tag{1}$$

(r and φ are the polar coordinates of the point of observation) everywhere except at the source — a "glowing line" parallel to the cylinder axis and having polar coordinates r' and 0; the primary wave produced by the source is

$$G^0 = \pi i H_0^{(1)}(k\rho), \quad \rho = \sqrt{r^2 - 2rr'\cos\varphi + r'^2}. \tag{2}$$

This means that on the ray $\varphi = 0$ the Green's function G is continuous, and its derivative $\partial G/\partial\varphi$ experiences a discontinuity

$$\left.\frac{\partial G}{\partial \varphi}\right|_{\varphi=+0} - \left.\frac{\partial G}{\partial \varphi}\right|_{\varphi=-0} = -4\pi r\delta(r - r'). \tag{3}$$

At infinity (as $r \to \infty$) the function G should satisfy the radiation condition, while on the cylinder surface it should satisfy the impedance boundary condition

$$\frac{\partial G}{\partial r} + ikgG = 0 \quad (\text{for} \quad r = a). \tag{4}$$

In acoustics, G is the velocity potential or the pressure, and 1/g is the normal impedance (total impedance). In electrodynamics (4) it is the Leontovich boundary condition and G is either the component of the magnetic field along the cylinder axis, or the corresponding component of the electric field; in the former case g = w, in the latter g = 1/w, where $w = \sqrt{\mu/\epsilon}$ is the wave resistance of the cylinder.

The function G satisfies the reciprocity relation

$$G(r, \varphi, r') = G(r', -\varphi, r) \tag{5}$$

and should satisfy the periodicity condition

$$G(r, \varphi + 2\pi j, r') = G(r, \varphi, r'), \; j = \pm 1, \pm 2, \ldots. \tag{6}$$

Along with the function G, it is advantageous to introduce the Green's function on the many-sheet plane* $\Gamma = \Gamma(r, \varphi, r')$, satisfying the relations (1) — (5), but obeying in lieu of the periodicity condition (6) the condition

$$\Gamma(r, \varphi, r') \to 0 \quad \text{for} \quad |\varphi| \to \infty. \tag{7}$$

The functions G and Γ are related by the formula

$$G(r, \varphi, r') = \sum_{j=-\infty}^{\infty} \Gamma(r, \varphi + 2\pi j, r'). \tag{8}$$

---

*The many-sheeted plane was first introduced for an analysis of diffraction around a cylinder in reference [10] (see also [11]).

The Green's function Γ is best represented not in the polar coordinates r and φ, but in the ray coordinates ξ and η, which are equivalent to this diffraction problem. These coordinates are introduced in the following fashion: assuming first that the observation point P is in the shadow region (Figure 1a, b) we draw from the point P a tangent PT to the circumference r = a of the cylinder, and we draw from the source P' a tangent P'T' and characterize the position of the point P with the aid of the coordinates

$$\xi = PT + \eta, \quad \eta = \overset{\frown}{T'T}, \tag{9}$$

and the position of the point P' with the aid of the difference

$$\xi' - \eta = P'T'.$$ (10

The coordinates ξ and η are uniquely related with the coordinates r and φ on a many-sheeted plane: thus, in Figure 1a, b, different ξ and η are set in correspondence with the poi P, but in Figure 1a the point P lies on the "zero" sheet of the plane $(-\pi < \varphi < \pi)$, while in Figure 1b it lies on the first order sheet $(-3\pi < \varphi < -\pi)$. The ray coordinates of the point P on the j-th order sheet (j = 1, 2, ...) are equal to $\xi + 2\pi a(j-1)$, $\eta + 2\pi a\,(j-1)$ (where ξ and η correspond to the zero-order sheet, see Figure 1a), while the ray coordinates of the point P on the - j-th sheet (j = 1, 2, ...) are $\xi + 2\pi a\,(j-1)$, $\eta + 2\pi a\,(j-1)$ (where ξ and η are its coordinates on the - 1-th sheet, (Figure 1b).

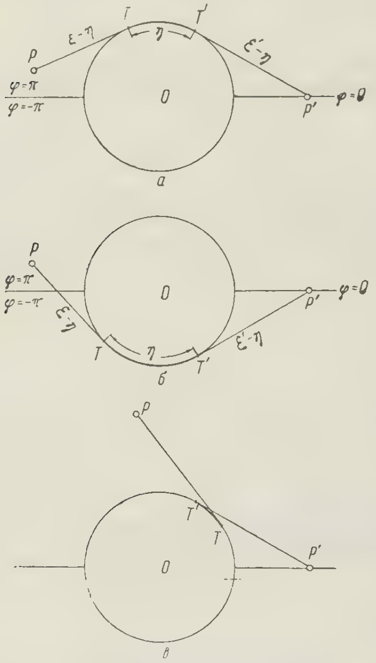


Figure 1. Ray coordinates for cylinder.

If the point P is located on the zeroth sheet, then it can fall in the eliminated region (see Figure 1c). In this case we introduce formally the ray coordinates ξ and η the same as in the shadow region, i.e., we draw tangents PT and P'T' and put

$$\xi - \eta = PT, \quad \xi' - \eta = P'T',$$ (11

assuming the coordinate η to be negative:

$$\eta = -\overset{\frown}{TT'}.$$ (12

It is physically obvious that in the eliminated region the coordinates ξ and η are suitable for the representation of the function Γ only for sufficientl small $|\eta|$, i.e., the boundary of the illuminated region, while away from this boundary the function Γ is represented as a sum of the primary wave (2) and the reflected waves, which obeys the laws of ge metric optics (for this wave the ray coordinates must be introduced in a different fashion).

However, the derivation of the laws of geometr optics from the wave equation (1) is hardly of any interest and the use of the ray coordinates (11) — (12) in the penumbral region, with $|\eta/a| \ll 1$, make it possible to effect a junction between the diffractio formulas and the formulas of geometric optics and thus obtain a complete solution of the problem unde consideration.

It is easy to show that ξ and η form an orthogonal coordinate system. The lines $\eta =$ = const are rays tangent to the periphery of the cylinder, while the lines ξ = const are involutes of this circle (see Figure 2), the Lamé coefficients being

$$h_\xi = 1,$$
$$h_\eta = \frac{\xi - \eta}{a}.$$ (13

The wave equation (1) assumes the following form in the ray coordinates

$$\frac{\partial^2\Gamma}{\partial\xi^2}+\frac{1}{\xi-\eta}\frac{\partial\Gamma}{\partial\xi}+$$
$$+\frac{a^2}{\xi-\eta}\frac{\partial}{\partial\eta}\left(\frac{1}{\xi-\eta}\frac{\partial\Gamma}{\partial\eta}\right)+$$
$$+k^2\Gamma=0. \tag{14}$$

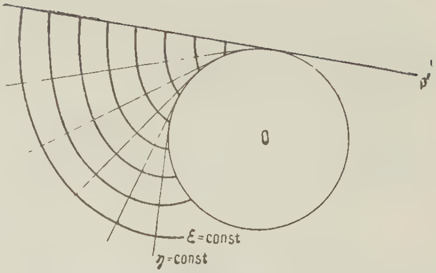


Figure 2. Rays (η = const) and involutes (ξ = const).

2. PARABOLIC EQUATION IN RAY COORDINATES

Proceeding to an approximate solution of (14) under the condition

$$M^2 \gg 1, \tag{15}$$

where

$$M=\left(\frac{ka}{2}\right)^{1/3}, \tag{16}$$

we first make the substitution

$$\Gamma=e^{ik\xi}W, \tag{17}$$

Separating the phase factor $e^{ik\xi}$ and introducing the function W, which is the slowly varying wave amplitude (generally speaking, complex). It satisfies the equation

$$\frac{\partial^2W}{\partial\xi^2}+\left(2ik+\frac{1}{\xi-\eta}\right)\frac{\partial W}{\partial\xi}+\frac{ik}{\xi-\eta}W+\frac{a^2}{\xi-\eta}\frac{\partial}{\partial\eta}\left(\frac{1}{\xi-\eta}\frac{\partial W}{\partial\eta}\right)=0, \tag{18}$$

which after introducing the dimensionless independent variables

$$X=M\frac{\xi}{a},\qquad Z=M\frac{\eta}{a} \tag{19}$$

assumes the form

$$4i\frac{\partial W}{\partial X}+\frac{2i}{X-Z}W+\frac{1}{X-Z}\frac{\partial}{\partial Z}\left(\frac{1}{X-Z}\frac{\partial W}{\partial Z}\right)+\frac{1}{M^2}\left(\frac{\partial^2W}{\partial X^2}+\frac{1}{X-Z}\frac{\partial W}{\partial X}\right)=0. \tag{20}$$

Neglecting in this equation terms of order $1/M^2$, we obtain an equation of the parabolic type

$$4i\frac{\partial W}{\partial X}+\frac{2i}{X-Z}W+\frac{1}{X-Z}\frac{\partial}{\partial Z}\left(\frac{1}{X-Z}\frac{\partial W}{\partial Z}\right)=0, \tag{21}$$

in which the third term, containing $\partial W/\partial Z$, takes into account the transverse diffusion of the wave amplitude, while terms of order $1/M^2$ in (20) correspond to "longitudinal diffusion," which which in this problem can be neglected (see [9]).

Equation (21) can be simplified. We introduce a new variable

$$Y=(X-Z)^2=M^2\left(\frac{\xi-\eta}{a}\right)^2;\quad Z=X-\sqrt{Y} \tag{22}$$

and regard W as the function of X and Y. Then Equation (21) becomes

$$i\frac{\partial W}{\partial X}+\frac{i}{2\sqrt{Y}}W+\frac{\partial^2W}{\partial Y^2}+2i\sqrt{Y}\frac{\partial W}{\partial Y}=0 \tag{23}$$

and after substituting

$$W=e^{-i\frac{2}{3}Y^{3/2}}W_1 \tag{24}$$

it is transformed into the Leontovich-Fok parabolic equation

$$i\frac{\partial W_1}{\partial X} + YW_1 + \frac{\partial^2 W_1}{\partial Y^2} = 0, \tag{25}$$

investigated in reference [2] . The boundary condition (4) assumes, on changing over to the function w_1, the form

$$\frac{\partial W_1}{\partial Y} + qW_1 = 0, \quad q = iMg \quad (\text{for } Y = 0), \tag{26}$$

since the height h of the point P over the surface of the cylinder

$$h = \sqrt{a^2 + (\xi - \eta)^2} - a \tag{27}$$

has the following value when $h \ll a$

$$h = \frac{(\xi - \eta)^2}{2a} = \frac{aY}{2M^2} = \frac{MY}{k}. \tag{28}$$

The general solution of (25) , satisfying the radiation condition and the boundary condition (26) , has the form

$$W_1 = \sum_{s=1}^{\infty} C_s e^{it_s X} w_1 (t_s - Y), \tag{29}$$

where C_s are unknown coefficients; w_1 is the Airy function corresponding to the outgoing wave, and t_s are the roots of the equation

$$w_1' (t_s) - qw_1 (t_s) = 0 \quad (s = 1, 2, \ldots), \tag{30}$$

which is investigated in detail in Section 7 of reference [12] .

The functions Γ, W and W_1 meaning also the coefficients C_s, depend on the position of the source, which we define in terms of the difference (10) and the corresponding dimensionless variable

$$Y' = M^2\left(\frac{\xi' - \eta}{a}\right)^2. \tag{31}$$

In order for W to be a slowly varying function of $\xi' - \eta$, it is necessary to separate in (17) in lieu of $e^{ik\xi}$, the phase factor

$$e^{ik(\xi+\xi'-\eta)} = e^{ik\sigma},$$

where

$$\sigma = \xi + \xi' - \eta \tag{32}$$

is the distance between the points P' and P, taken along the minimal path (diffraction ray) consisting of the straight-line segments P'T' and TP and the are T'T. Recognizing further that the variables ξ' and ξ, as well as Y' and Y, should enter into the function Γ in a symmetrical fashion, we arrive at the expression

$$\Gamma = e^{i\left(k\sigma - \frac{2}{3} Y^{3/2} - \frac{2}{3} Y'^{3/2}\right)} \sum_{s=1}^{\infty} c_s e^{it_s X} w_1 (t_s - Y) w_1 (t_s - Y'), \tag{33}$$

where X is determined by a new formula

$$X = M\frac{\sigma}{a} = M\frac{\xi + \xi' - \eta}{a}, \tag{34}$$

in which ξ and ξ' are symmetrical. When $\xi' = \eta$, when the point P' lies on the surface of the cylinder, formula (34) goes into the initial definition (19) of X; the transition from (29) to (33) signifies that the coefficient C_s contains the factor $\exp\{it_s M(\xi' - \eta)/a\}$.

3. DETERMINATION OF THE COEFFICIENTS c_s (BOTH POINTS NEAR THE CYLINDER SURFACE)

The coefficients c_s in (33) can depend only on the parameters q and M. They can be determined from the following considerations (cf [2] and [13]). Let the source P' and the point of observation P lie near the surface of the cylinder, so that their heights h' = r' - a amd h = r - a satisfy the conditions

$$h' \ll a, \quad h \ll a. \tag{35}$$

Then the variables Y and Y' can be replaced in (33) by Y and Y' where, according to (28)

$$y = \frac{kh}{M}, \qquad y' = \frac{kh'}{M}, \tag{36}$$

and we can put in addition

$$e^{i\left(k\sigma - \frac{2}{3}Y^{3/2} - \frac{2}{3}Y'^{3/2}\right)} = e^{ika|\varphi|} = e^{ikl}, \tag{37}$$

where $l = a\,|\varphi|$ is the distance between the Points P' and P along the surface of the cylinder. The last relation follows from Figure 3, where

$$\left.\begin{aligned} \angle P'OP = |\varphi| = \frac{\eta}{a} + \Delta\varphi + \Delta\varphi', \\ \Delta\varphi = \operatorname{arc\,tan}\frac{\xi-\eta}{a} = \frac{\xi-\eta}{a} - \frac{1}{3}\left(\frac{\xi-\eta}{a}\right)^3 + \dots, \end{aligned}\right\} \tag{38}$$

so that for small $\Delta\varphi$ and $\Delta\varphi'$

$$\left.\begin{aligned} ka\Delta\varphi = k(\xi-\eta) - \frac{2}{3}Y^{3/2}, \quad ka\Delta\varphi' = k(\xi'-\eta) - \frac{2}{3}Y'^{3/2}, \\ ka\,|\varphi| = k\sigma - \frac{2}{3}Y^{3/2} - \frac{2}{3}Y'^{3/2}. \end{aligned}\right\} \tag{39}$$

At finite Y and Y' we can also replace X by

$$x = M\,|\varphi| = M\frac{l}{a}, \tag{40}$$

and the formula (33) becomes

$$\Gamma = e^{ikl}\sum_{s=1}^{\infty} c_s e^{it_s x}\, w_1(t_s - y)\, w_1(t_s - y'). \tag{41}$$

The coefficients c_s must be chosen such that in the eliminated region the following reflection formula

$$\Gamma = \sqrt{\frac{2\pi}{kl}}\, e^{i\left(kl + \frac{\pi}{4}\right)}\left[e^{\frac{ik(h-h')^2}{2l}} + e^{\frac{ik(h+h')^2}{2l}}\,\frac{h+h'-gl}{h+h'+gl}\right], \tag{42}$$

corresponding to a plane reflecting surface and sufficiently gently sloping rays is satisfied. The first term in the square bracket of (42) gives the primary wave (2)

$$G^0 = \sqrt{\frac{2\pi}{k\rho}}\, e^{i\left(k\rho + \frac{\pi}{4}\right)}, \quad \rho = \sqrt{l^2 + (h-h')^2} \simeq l + \frac{(h-h')^2}{2l}, \tag{43}$$

while the second term gives the reflected wave, coming from a virtual source and having a reflection coefficient

$$\frac{\cos\chi - g}{\cos\chi + g} \simeq \frac{\frac{h+h'}{l} - g}{\frac{h+h'}{l} + g}, \tag{44}$$

and where χ is the angle of incidence (close to $\pi/2$)

If we put

$$\Gamma = \frac{2\pi i}{M}\, e^{ikl}\Psi(x, y, y', q), \tag{45}$$

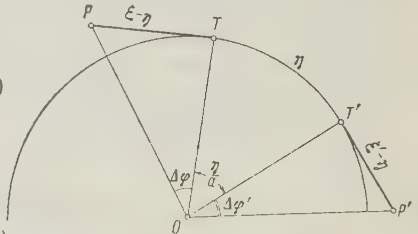


Figure 3. Angles $\Delta\varphi$ and $\Delta\varphi'$ in the case when both points are located near the cylinder.

where the function ψ is determined by the contour integral (the contour C includes all the points t_s in the positive direction)

$$= \frac{1}{4\pi}\int_C e^{itx} w_1(t-y')\left[w_2(t-y) - \frac{w_2'(t) - qw_2(t)}{w_1'(t) - qw_1(t)}\, w_1(t-y)\right] dt \tag{46}$$

and the series of the residues

$$\Psi(x, y, y', q) = \sum_{s=1}^{\infty} \frac{e^{it_s x}}{t_s - q^2} \frac{w_1(t_s - y)}{w_1(t_s)} \frac{w_1(t_s - y')}{w_1(t_s)}, \tag{47}$$

then by virtue of the formula (cf [13] , Section 5)

$$\Psi = \frac{e^{-i\frac{\pi}{4}}}{2\sqrt{\pi x}} \left[e^{i\frac{(y-y')^2}{4x}} + e^{i\frac{(y+y')^2}{4x}} \frac{y + y' + 2iqx}{y + y' - 2iqx} \right] (x \to 0) \tag{48}$$

we actually obtain the reflection formula (42) . It is difficult to prove the uniqueness of the solution (46) (cf. [2]) ; we can only refer to part II of that paper [19] , where a more rigorous derivation of (45) and (46) is given.

We note that in writing out the integral (46) , it is assumed that y' ≥ y. If, to the contrary y' < y, we must interchange y and y'. The function ψ is connected with the attenuation factor V which is used in connection with the theory of propagation of radio waves around the earth's surface, and is partially tabulated in [14] and [15] , by means of the relation

$$V(x, y, y', q) = 2\sqrt{\pi x}\, e^{i\frac{\pi}{4}} \Psi(x, y, y', q). \tag{49}$$

are also important; they are defined in Section 3 of reference [14] .

As already noted, formula (42) corresponds to gently sloping rays and should be supplemented with a formula suitable for greater heights h and h'. On the other hand, as h and h' are decreased or as l is increased, formula (42) becomes useless since it is necessary to take into account the curvature of the reflecting surface (and in particular the additional divergence of the reflected rays, see [14] , Section 5) , while at small $|g|$ it is necessary to use the Weyl — Van der Pol formula, let alone the fact that diffraction phenomena come into play near the boundary of the illuminated region.

Comparing (45) and (47) with (41) , we get

$$c_s = \frac{2\pi i}{M} \frac{1}{(t_s - q^2)\, w_1^2(t_s)}. \tag{50}$$

Inasmuch as the coefficients c_s cannot depend on Y and Y', the general formula (33) assumes the form

$$\Gamma = \frac{2\pi i}{M} e^{i\left(k\sigma - \frac{2}{3} Y^{3/2} - \frac{2}{3} Y'^{3/2}\right)} \Psi(X, Y, Y', q). \tag{51}$$

This expression, which is suitable for all distances from the source and from the point of observation to the cylinder, is the main result of the present paper. We note that it is obtained from formula (45) , which is valid only under conditions (35) , by replacing the factor e^{ikl} by more complicated one, and by substituting in the function ψ new variables X, Y, and Y', connected with the ray coordinates, in lieu of the variables x, y, and y'.

It will be shown later that (51) agrees with results obtained in other papers in which the problem has a more specialized formulation, and converges to the formulas of geometric optics in the illuminated region.

4. ONE POINT NEAR THE CYLINDER AND THE OTHER AWAY FROM IT

We now remove conditions (35) . In this section we assume that h ≪ a, but the height h' can be as large as desired, for example h' ~ a or h' ≫ a. Then Y' will be large, and we can employ the relation

$$\Psi(X, Y, Y', q) = \frac{1}{2\sqrt{\pi}\, Y'^{1/4}} e^{i\left(\frac{2}{3} Y'^{3/2} - \frac{\pi}{4}\right)} V_1(z, Y, q), \tag{52}$$

where

$$z = X - \sqrt{Y'} = M\frac{\xi}{a}; \tag{53}$$

$$V_1(z, y, q) = \frac{i}{2\sqrt{\pi}}\int\limits_C e^{itz}\left[w_2(t-y) - \frac{w_2'(t) - qw_2(t)}{w_1'(t) - qw_1(t)}\,w_1(t-y)\right]dt. \tag{54}$$

Formula (51) is rewritten in the form

$$\Gamma = \sqrt{\frac{2\pi}{k(\xi' - \eta)}}\,e^{i\left[k(\xi'-\eta)+\frac{\pi}{4}\right]}\,e^{i\left(k\xi - \frac{2}{3}Y^{3/2}\right)}V_1(z, Y, q). \tag{55}$$

This last expression breaks up into two factors; the factor

$$G^0 = \sqrt{\frac{2\pi}{k(\xi' - \eta)}}\,e^{i\left[k(\xi'-\eta)+\frac{\pi}{4}\right]} \tag{56}$$

corresponds to free propagation of the primary cylindrical wave (2) from the source P' to the point of tangency T'; the second factor corresponds to diffraction propagation along the path P'TP. If the primary wave is plane

$$G^0 = e^{-ikr\cos\varphi}, \tag{57}$$

then formula (55) simplifies to

$$\Gamma = e^{i\left(k\xi - \frac{2}{3}Y^{3/2}\right)}V_1(z, Y, q). \tag{58}$$

The meaning of ξ and η is illustrated in Figure 4. We denote

$$\angle POT' = \psi = \frac{\eta}{a} + \Delta\varphi; \tag{59}$$

and then by virtue of the first relation of (39)

$$ka\psi = k\xi - \frac{2}{3}Y^{3/2} \tag{60}$$

and in lieu of (57) we can write, by replacing Y with small y,

$$\Gamma = e^{ika\psi}V_1(z, y, q), \tag{61}$$

where z = Mψ.

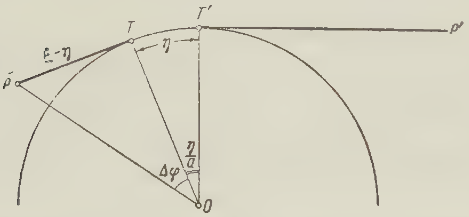


Figure 4. Coordinates ξ, η, and the angle $\Delta\varphi$ for a plane wave (or for a remote source P').

The last expression agrees with the formulas derived in [16], and [17]. In the illuminated region, z and ψ are negative, and for large negative z the function V_1 satisfies the asymptotic formula

$$V_1(z, y, q) = e^{-\frac{i}{3}z^3}\left(e^{izy} + \frac{z - iq}{z + iq}\,e^{-izy}\right). \tag{62}$$

It follows from this formula that (61) assures convergence to the geometrical-optics relations: the first term in the parentheses of (62) gives the incident wave, and the second the reflected one, and $\pi/2 + \psi$ is the angle of incidence.

In the illuminated region ($\psi < 0$, $z < 0$) we can employ instead of (61) the following interpolation formula (cf. [16])

$$\Gamma = e^{ika\sin\psi}e^{\frac{i}{3}z^3}V_1(z, y, q), \quad z = M\sin\psi, \tag{63}$$

which for finite z and small ψ almost coincides with (61), and which for large negative z, by virtue of (62), goes into the reflection formula

$$\Gamma = e^{-ik(a+h)\cos\varphi} + e^{-ik(a-h)\cos\varphi}\,\frac{\cos\varphi - g}{\cos\varphi + g}, \tag{64}$$

which is suitable for any angle of incidence $\varphi = \frac{\pi}{2} + \psi$.

It is easy to verify that for an incident cylindrical wave (56), produced by a remote source, formula (55) also provides convergence to geometrical optics for finite Y and $z < 0$. By virtue of the reciprocity relation (5), the foregoing formulas are suitable also for the case when the source is near the cylinder and the point of observation far from it

We note that the "reflection formula" (62) for the function V_1 can be made more precise in such a manner, that it takes account of the influence of the curvature of the reflecting surface on the reflected wave (see [4], Section 5, or [14], Section 5); at sufficiently large negative z and finite y, however, this influence can be neglected.

5. BOTH POINTS FAR FROM CYLINDER

Let us consider first the case when the incident wave is plane, and let us regard Y in formula (58) as a large quantity (finite and small values of Y have been considered above). We introduce the variable

$$\zeta = z - \sqrt{Y} = M \frac{\eta}{a} \tag{65}$$

and use the relations (see [14], Section 3)

$$\left.\begin{aligned} V_1(z, Y, q) &= \frac{1}{Y^{1/4}} e^{i\frac{2}{3} Y^{3/2}} V_{11}(\zeta, Y^{1/4}, q) \quad \text{for} \quad \zeta > 0, \\ V_1(z, Y, q) &= e^{-\frac{i}{3} z^3 + izY} + \frac{1}{Y^{1/4}} e^{i\frac{2}{3} Y^{3/2}} V_{11}(\zeta, Y^{1/4}, q) \quad \text{for} \quad \zeta < 0, \end{aligned}\right\} \tag{66}$$

where

$$V_{11}(\zeta, \mu, q) = V^{(0)}(\zeta, \mu) + V^{(1)}(\zeta, q) + V^{(2)}(\zeta, q). \tag{67}$$

The functions $V^{(0)}$, $V^{(1)}$, and $V^{(2)}$ are defined as follows

$$\left.\begin{aligned} V^{(0)}(\zeta, \mu) &= \mu \frac{e^{-i\left(\mu^2\zeta^2 + \frac{\pi}{4}\right)}}{\sqrt{\pi}} \int_{\mu\zeta}^{\infty} e^{is^2}\, ds \quad \text{for} \quad \zeta > 0, \\ V^{(0)}(\zeta, \mu) &= -\mu \frac{e^{-i\left(\mu^2\zeta^2 + \frac{\pi}{4}\right)}}{\sqrt{\pi}} \int_{-\infty}^{\mu\zeta} e^{is^2}\, ds \quad \text{for} \quad \zeta < 0 \end{aligned}\right\} \tag{68}$$

and

$$\left.\begin{aligned} V^{(1)}(\zeta, q) &= -\frac{e^{-i\frac{\pi}{12}}}{\sqrt{\pi}} \int_0^{\infty} e^{-t\zeta e^{i\frac{\pi}{6}}} \frac{v'(t) + e^{-i\frac{\pi}{3}} q v(t)}{w_2'(t) + e^{-i\frac{\pi}{3}} q w_2(t)}\, dt, \\ V^{(2)}(\zeta, q) &= -\frac{e^{i\frac{\pi}{4}}}{\sqrt{\pi}} \int_0^{\infty} e^{it\zeta} \frac{v'(t) - q v(t)}{w_1'(t) - q w_1(t)}\, dt. \end{aligned}\right\} \tag{69}$$

In (66) we neglect terms of order $1/\mu^2 = 1/\sqrt{Y}$ compared with terms of order unity (see [18], Section 5), and therefore for finite and large ζ we must use in lieu of the Fresnel integrals (68) the first term of their asymptotic expansions, and write

$$V^{(0)}(\zeta, \mu) = V^{(0)}(\zeta, \infty) = -\frac{e^{i\frac{3\pi}{4}}}{2\sqrt{\pi}\,\zeta}. \tag{70}$$

Then the function V_{11} depends only on two variables and is transformed to

$$V_{11}(\zeta, q) = V_{11}(\zeta, \infty, q). \tag{71}$$

The dependence of V_{11} on variable μ manifests itself only at small values of ζ.

By virtue of (66), formula (58) yields

$$\left.\begin{aligned} \Gamma &= M \sqrt{\frac{2}{k(\xi - \eta)}} e^{ik\xi} V_{11}(\zeta, Y^{1/4}, q) \quad \text{for} \quad \zeta > 0, \\ \Gamma &= \Gamma^0 + M \sqrt{\frac{2}{k(\xi - \eta)}} e^{ik\xi} V_{11}(\zeta, Y^{1/4}, q) \quad \text{for} \quad \zeta < 0, \end{aligned}\right\} \tag{72}$$

where the term

$$\Gamma^0 = e^{i\left(k\xi - \frac{2}{3} Y^{3/2} - \frac{1}{3} z^3 + zY\right)} = e^{i\left(k\xi - \frac{1}{3}\zeta^3 - \zeta^2\sqrt{Y}\right)} \tag{73}$$

yields for small η/a the incident wave (57). In fact

$$-r\cos\varphi = (\xi-\eta)\cos\frac{\eta}{a} + a\sin\frac{\eta}{a}, \tag{74}$$

and putting

$$\cos\frac{\eta}{a} = 1 - \frac{\eta^2}{2a^2}, \qquad \sin\frac{\eta}{a} = \frac{\eta}{a} - \frac{\eta^3}{6a^3}, \tag{75}$$

we obtain

$$-kr\cos\varphi = k\xi - \frac{1}{3}\zeta^3 - \zeta\sqrt{Y}. \tag{76}$$

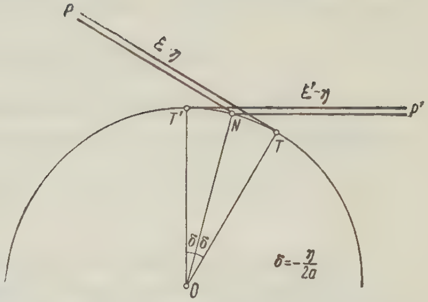


Figure 5. Reflection of rays and ray coordinates (for remote points P' and P.

Actually there is no point in distinguishing between Γ^0 and G^0 so that Γ^0 should be taken to mean merely the incident wave (57). The reflected wave is

$$\Gamma^1 = M\sqrt{\frac{2}{k(\xi-\eta)}}\, e^{ik\xi} V_{11}(\zeta, Y^{1/4}, q) \ (\text{for } \zeta < 0). \tag{77}$$

For large negative ζ, we have the following asymptotic expression for V_{11}

$$V_{11} = \frac{\sqrt{-\zeta}}{2} e^{-\frac{i}{12}\zeta^3} \frac{-\frac{\zeta}{2} + iq}{-\frac{\zeta}{2} - iq}, \tag{78}$$

which ensures convergence of (77) to the geometrical-optics case, according to which we should have

$$\Gamma^1 = \frac{M^{3/2}}{\sqrt{k(\xi-\eta)}} e^{ik(\xi-\eta-2a\sin\delta)} \sqrt{\sin\delta}\, \frac{\sin\delta - g}{\sin\delta + g}, \tag{79}$$

where $\pi/2-\delta$ is the angle of reflection of the given ray NP ($\delta = -\eta/2a$, see Figure 5). Instead of the two formulas (77) and (79) we can write for $\zeta < 0$ (cf. [16]) a single interpolation formula

$$\Gamma^1 = M\sqrt{\frac{2}{k(\xi-\eta)}}\, e^{ik\left(\xi-\eta+2a\sin\frac{\eta}{2a}\right)} e^{-\frac{i}{12}\zeta^3} V_{11}(\zeta, Y^{1/4}, q), \tag{80}$$

where $\pi/2 = 2M \sin \eta/2a$; this interpolation formula almost coincides with (77) for finite ζ and with (79) when $-\zeta \gg 1$. When $q = 0$ and $q = \infty$, and also when we replace $V_{11}(\zeta, Y^{1/4}, q)$ by $V_{11}(\zeta, q)$ formulas (72), (77), and (80) become equivalent to the expressions derived by A.S. Goryainov [16] using a different method.

If a cylindrical wave (56) produced by a remote source is incident on the cylinder, and Y' is comparable with Y, we must start out not from (55), but directly from the general formula (51). Using the connections between V_{11} and V [see formula (49)], established in Fok's paper [18] and cited in [14], we obtain

$$\Gamma = M\sqrt{\frac{4\pi}{k^2(\xi'-\eta)(\xi-\eta)}}\, e^{i\left(k\sigma+\frac{\pi}{4}\right)} V_{11}(\zeta, \mu, q), \tag{81}$$

where

$$\zeta = X - \sqrt{Y} - \sqrt{Y'} = M\frac{\eta}{a} > 0, \quad \mu^2 = \frac{\sqrt{YY'}}{\sqrt{Y} + \sqrt{Y'}}. \tag{82}$$

The same formula (81) is suitable also for the reflected wave Γ^1 with finite negative ζ. We can use Γ^1 also the interpolation formula

$$\Gamma^1 = M\sqrt{\frac{4\pi}{k^2(\xi'-\eta)(\xi-\eta)}}\, e^{i\left(k\sigma+\frac{\pi}{4}+2a\sin\frac{\eta}{2a}\right)} e^{-\frac{i}{12}\zeta^3} V_{11}(\zeta, \mu, q), \tag{83}$$

where $\zeta = 2M \sin \eta/2a < 0$, which provides an automatic transition to the geometrical optics case provided ξ and ξ' are large compared with $|\eta|$, so that we can assume in Figure 5

$$P'T' \parallel P'N \text{ and } TP \parallel NP. \tag{84}$$

The formulas of this section can be readily interpreted from the point of view of diffraction rays. A deeper interpretation of the diffraction phenomena described by these formulas is based on the fact that the function $V^{(0)}$ in (67) corresponds to diffraction on a semiinfinite plane screen, which screens the incident wave in the same way as the given cylinder, and the sum $V^{(1)} + V^{(2)}$ gives a diffraction "background" — a correction to $V^{(0)}$ dependent on the curvature and material of the cylinder (cf. [18] and [14]). The resolution of V_{11} into $V^{(0)}$ and $V^{(1)} + V^{(2)}$ is physically connected with the presence of two zones of "effective transverse diffusion" in the given problem (see [9] .

CONCLUSION

The main result of the present paper is expression (51) which converges to the formulas of geometric optics and gives a general solution to the problem of diffraction on a circular cylinder of large radius. This result is evidence of the effectiveness of the parabolic equation written in ray coordinates.

REFERENCES

1. M.A. Leontovich, Izv. AN SSSR, ser fiz, 1944, 8, 1, 16.
2. M.A. Leontovich and V.A. Fok, ZhETF, 1946, 16, 7, 557; Sb. Issledovaniya po rasprostraneniyu radiovoln, (Collection, Research on Radio Wave Propagation), edited by B.A. Vvedenskiy, AN SSSR, 1948, pp 13 — 39.
3. V.A. Fok, Izv. AN SSSR, ser fiz, 1946, 10, 2, 171.
4. V.A. Fok, Uspekhi fiz. nauk, 1948, 36, 3, 308.
5. G.D. Malyuzhinets, Paper delivered at All-Union Conference on Problems of Electric Oscillations and Waves in Gor'kii, 1946.
6. G.D. Malyuzhinets, Paper delivered at Fourth-All-Union Acoustic Conference, Moscow, 1958, abstracts of papers, Lithographed edition of the organizing committee of the conference, Part II, 1958, p. 3.
7. G.D. Malyuzhinets, Paper delivered at Third International Acoustic Conference, Stuttgart, 1959.
8. G.D. Malyuzhinets, Paper delivered at Symposium on Wave Diffraction, Odessa, 1960 Synopsis of paper, AN SSSR, 1960, p. 7.
9. G.D. Malyuzhinets, Uspekhi fiz. nauk, 1959, 69, 2, 321.
10. G.D. Malyuzhinets, Certain Generalizations of the Method of Reflection in the Theory of Diffraction. Author's abstract of doctoral dissertation, published by AN SSSR, 1950.
11. F.G. Friedlander, Communications in Pure and Applied Mathematics, 1954, 7, 4, 705.
12. V.A. Fok, Diffraktsiya radiovoln vokrug zemnoy poverkhnosti, (Diffraction of Radio Waves Around the Earth's Surface), Izd. AN SSSR, 1946.
13. V.A. Fok, Izv. AN SSSR, ser. fiz, 1950, 14, 1, 70.
14. P.A. Azrilyant, M.G. Belkina, Chislennye rezul'taty teorii difractsii radiovoln vokrug zemn poverkhnosti, (Numerical results for the theory of radio wave diffraction around the earth's surface) Izd. AN SSSR, 1957.
15. M.G. Belkina, Tablitsy dlya vycheslenniya elektromagnitnogo polya v oblasti teni dlya razlichnykh pochv, (Tables for the calculation of the electromagnetic field in the shadow region for different soils), Sovetskoye radio, 1949.

16. A.S. Goryainov, Radiotekhnika i elektronika, 1958, 3, 5, 603, (see also Dokl. AN SSSR , 1956, 109, 3, 477).
17. J.R. Wait and A.M. Conde, IRE Transactions, 1958, AP-6, 4, 348.
18. V.A. Fok, Uspekhi fiz. nauk, 1950, 43, 4, 587.
19. L.A. Vaynshteyn and G.D. Malyuzhinets, Radiotekhnika i elektronika, 1961, 6, 9.

Acoustic Institute Academy of Sciences USSR,
Institute of Physics Problems Academy of Sciences, USSR.

Submitted to the editors 18 January 1961

EXCITATION OF WAVES ON AN IMPEDANCE PLANE

M.D. Khaskind

An investigation is made of the two-dimensional field of E waves above an impedance plane; this field is excited by a system of sources contained in an ideally conducting cylindrical surface with slots cut parallel to the generatrices. To determine the wave field, a generalized Green-Kirchhoff formula is derived. A complex wave amplitude is then introduced and used to express the energy characteristics of the total field and the surface-wave field.

1. GENERALIZED GREEN-KIRCHHOFF FORMULA AND ENERGY RELATIONS

1. Fundamental Equations

We consider an impedance plane z = 0 and investigate the two-dimensional electromagnetic field in the upper half-space outside a cylindrical surface with transverse contour L and a generatrix parallel to the x-axis (Figure 1). The time dependence is accounted for by an exponential factor exp $i\omega t$, and the rationalized practical system of units is used.

On the impedance plane we have the condition

$$E_y = i\rho_0 Z H_x \quad \text{for} \quad z = 0 \left(\rho_0 = \left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2} = 120\pi\,\mathbf{ohm}\right), \tag{1}$$

in which Z is considered positive for simplicity. In particular, in the case of a comblike structure we have Z = tan kl, where l is the height of the comb and $k = \omega(\epsilon_0\mu_0)^{1/2}$ is the wave number in free space.

We determine the electromagnetic field above the plane z = 0 with the aid of the relations

$$\mathbf{E} = -ik\rho_0 \,\mathrm{curl}\, \mathbf{\Pi}_m, \quad \mathbf{H}_m = \mathrm{grad\ div}\ \mathbf{\Pi}_m + k^2 \mathbf{\Pi}_m,$$

where the Hertz magnetic vector Π_m is directed along the x-axis by putting $\Pi_{mx} = \varphi(y, z)$ and $\Pi_{my} = \Pi_{mz} = 0$. As a result of this we obtain (x^0 is the unit vector along the x-axis)

$$\mathbf{E} = -ik\rho_0 (\mathrm{grad}\ \varphi \times \mathbf{x}^0), \quad E_x = 0, \quad E_y = -ik\rho_0 \frac{\partial\varphi}{\partial z}, \quad E_z = ik\rho_0 \frac{\partial\varphi}{\partial y}, \tag{2}$$
$$H_x = k^2\varphi, \quad H_y = H_z = 0.$$

The scalar function $\varphi(y, z)$ satisfies the wave equation and, according to (1) and (2), the boundary equation

$$\frac{\partial\varphi}{\partial z} + p\varphi = 0 \text{ for } z = 0 \ (p = kZ). \tag{3}$$

We assume further that the system of sources is located inside a perfectly conducting cylindrical surface, in which slots are cut along the generatrix. We then have on this surface

$$\frac{\partial\varphi}{\partial n} = \frac{1}{ik\rho_0} E_l \text{ on } L, \tag{4}$$

where the tangential component E_l of the electric field intensity vanishes on the perfectly-conducting parts, while on the slots E_l is either a specified function or a function to be determined by "matching" the internal and external fields. In the case of concentrated slots with specified voltages V_s applied to their edges ($s = 0, 1, \ldots m$), condition (4) can be expressed in δ-function form

$$\frac{\partial\varphi}{\partial n} = \frac{1}{ik\rho_0} \sum_{s=0}^{m} V_s \delta(l - l_s) \text{ on } L. \tag{5}$$

Each concentrated slot excites the same field as a linear magnetic current $I_{ms} = 2V_s$. The form of the remaining part of the perfectly conducting contour L manifests itself in the fact that electric currents of density H_x are produced on its outer side.

Let us present various forms for the source function of cylindrical waves over the impedance plane $z = 0$. For this purpose we represent the source function $G(y, z, a, \zeta)$, located at the point $C(\eta, \xi)$, in the form of the sum

$$\begin{gathered} G = G_1 + G_2 + G_3, \\ G_1 = H_0^{(2)}\ (kr_1), \\ G_2 = H_0^{(2)}\ (kr_2), \\ (r_1 = ((y - \eta)^2 + (z - \zeta)^2)^{1/2}, \\ r_2 = ((y - \eta)^2 + (z + \zeta)^2)^{1/2},) \end{gathered} \tag{6}$$

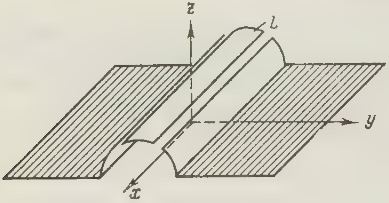


Figure 1

Here G_1 is the source function for cylindrical waves in free space; the function G_2 determines the specularly reflected field produced in the presence of a perfectly conducting plane $z = 0$, and the function G_3 is due to the finite value of the impedance of the plane $z = 0$. As shown in reference [1], we have for the function G_3 two identical expressions, which differ outwardly from each other

$$\begin{gathered} G_3 = \frac{p}{hi}\left(e^{-ihy}\int_{\infty}^{y} e^{-ih\eta_0} F d\eta_0 - e^{-ihy}\int_{-\infty}^{y} e^{ih\eta_0} F d\eta_0\right) \\ \left(F = \left(\left(\frac{\partial}{\partial z} - p\right) H_0^{(2)}(kr_2)\right)_{y=\eta_0}\right), \end{gathered} \tag{7}$$

$$\begin{gathered} G_3 = -\,2pe^{-pz}\int_{-\infty}^{z} e^{p\tau} H_0^{(2)}(kr_2')\,d\tau + \frac{4p}{h} e^{-p(z+\zeta)-ih|y-\eta|} \\ (r_2' = (r_2)_{z=\tau}, \quad h^2 = k^2 + p^2), \end{gathered} \tag{8}$$

where the integration in (7) is along half-lines parallel to the y-axis and located above the point, $C^*(\eta, -\xi)$, which the linear path of integration $(-\infty, z)$ in (8) does not contain, the second term in (8) determines the excited surface-wave field. We note also that the expression for G_3 can be represented with the aid of a single contour integral in the complex τ plane. For this purpose it is sufficient to introduce temporarily a slight amount of absorption, i.e., to assume $\mathrm{Im}\, k < 0$, $\mathrm{Im}\, p < 0$, $\mathrm{Re}\, p > 0$; we then have [2, 3]

$$G_3 = -2pe^{-pz} \int_{-i\infty}^{z} e^{p\tau} H_0^{(2)}(kr_2')\, d\tau. \tag{9}$$

If we now use the expansion

$$H_0^{(2)}(kr_2') = \frac{i}{\pi} \int_{-\infty}^{\infty} e^{-iw(y-\eta)-u(\tau+\zeta)} \frac{dw}{u} \quad (u = (w^2 - k^2)^{1/2}, \quad \mathrm{Re}\, u > 0), \tag{10}$$

we can transform (9) into

$$G_3 = \frac{2pi}{\pi} \int_{-\infty}^{\infty} \frac{e^{-iw(y-\eta)-u(z+\zeta)}}{u(u-p)}\, dw. \tag{11}$$

In the limit as $\mathrm{Im}\, k = 0$, the branch point $w = \pm k$ becomes real, and accordingly the integration path in (11) must pass around the branch points $w = \pm k$ and the poles $w = \pm h$ in the manner indicated in Figure 2.

We transform the contour integration in (11) into integration along the imaginary axis in the w-plane. Making a further suitable change in variables, we obtain

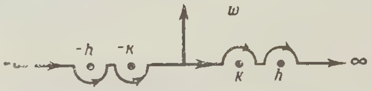


Figure 2

$$G_3 = \frac{2pi}{\pi} \int_{-\infty}^{\infty} \frac{e^{-iv(z+\zeta)-m|y-\eta|}}{m(iv-p)}\, dv + \frac{4p}{h} e^{-p(z+\zeta)-ih|y-\eta|} \tag{12}$$

$$(m = (v^2 - k^2)^{1/2}, \quad \mathrm{Re}\, m > 0).$$

Relation (12) follows directly from (8). In fact, by substituting in (8) the expansion

$$H_0^{(2)}(kr_2') = \frac{i}{\pi} \int_{-\infty}^{\infty} e^{-iv(z+\zeta)-m|y-\eta|} \frac{dv}{m}, \tag{13}$$

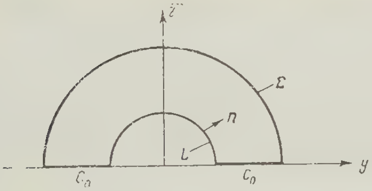


Figure 3.

we arrive directly at (12) and verify thereupon the identity o (8) and (9).

Let us proceed now to establish the generalized Green-Kirchhoff formula. Applying for this purpose, Green's formula to the functions φ and G, we obtain

$$\int_{L+\Sigma+C_0} \left(G \frac{\partial \varphi}{\partial n} - \varphi \frac{\partial G}{\partial n}\right) dl + 4i\varphi(y, z) = 0, \tag{14}$$

where Σ is an arbitrary contour in the upper half-plane, encompassing the contour L and contiguous in part with the y-axis; C_0 is the free part of this axis between L and Σ (Figure 3) When $z = 0$, the functions φ and G satisfy the condition (3) and consequently the integral with respect to C_0 in (14) vanishes automatically. To transform Σ into an infinitely remote contour, we bring into play a small absorption ($\mathrm{Im}\, k < 0$), which ensures the damping of the field at infinite distances. As a result we arrive at the generalized Green-Kirchoff formula

$$\varphi = \frac{i}{4} \int_{L} \left(G \frac{\partial \varphi}{\partial n} - \varphi \frac{\partial G}{\partial n}\right) dl, \tag{15}$$

indicated in [1] and used in [4, 5] in a somewhat different form ($k = ik_1$).

Let us note certain general properties that follow from (15). First, it is easy to separate the surface-wave field with the aid of (15). For this purpose we consider only the second term in (8); we then obtain

$$\varphi = \frac{ip}{h} A(\pm h, p)\, e^{-pz \mp ihy} \quad \text{for } y \to \pm\infty, \tag{16}$$

$$A(\pm h, p) = \int_L e^{\pm ihy - pz} \left[\frac{\partial\varphi}{\partial n} - \varphi\,(\pm ih\cos(n, y) - p\cos(n, z))\right] dl, \tag{17}$$

where the functions $A(\pm h, p)$ determine, accurate to a factor iph^{-1}, the complex amplitudes of the surface-waves that diverge on both sides of the contour L. We next represent (15), in accordance with (6), in the form of a sum

$$\varphi = \varphi_1 + \varphi_2 + \varphi_3, \quad \varphi_s = \frac{i}{4}\int_L \left(G_s \frac{\partial\varphi}{\partial n} - \varphi\frac{\partial G_s}{\partial n}\right) dl \quad (s = 1, 2, 3). \tag{18}$$

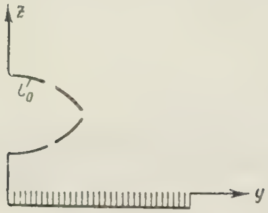


Figure 4

If the values of φ and of its normal derivative are specified on the contour L, then relation (18) can be interpreted exactly as (6), i.e., the function φ_1 determines the field in free space, the function φ_2 characterizes the specularly reflected field which occurs in the presence of a perfectly conducting surface $z = 0$, and finally the function φ_3 takes into account the finite value of the impedance of the plane $z = 0$. On the basis of (10) and (11), we express the functions φ_2 and φ_3 in terms of the complex amplitude $A(w, u)$ in the following manner:

$$\varphi_2 = -\frac{1}{4\pi}\int_{-\infty}^{\infty} e^{-iwy - uz} A(w, u)\frac{dw}{u}, \qquad \varphi_3 = -\frac{p}{2\pi}\int_{-\infty}^{\infty} e^{-iwy - uz}\frac{A(w, u)}{u(u - p)}\, dw. \tag{19}$$

In addition, as follows from (18) and (6), the functions φ_1 and φ_2 have asymptotic forms that are characteristic of cylindrical waves;

$$\varphi_1 = \frac{iA(k\cos\theta, -ik\sin\theta)}{(8\pi kr)^{1/2}} e^{-i\left(kr - \frac{\pi}{4}\right)}, \quad \varphi_2 = \frac{iA(k\cos\theta, ik\sin\theta)}{(8\pi kr)^{1/2}} e^{-i\left(kr - \frac{\pi}{4}\right)}$$
$$\left(r = (y^2 + z^2)^{1/2}, \quad \theta = \operatorname{arc\,tan}\frac{y}{z}\right). \tag{20}$$

Let us note another case of unilateral propagation of waves. Let a cylindrical surface with transverse contour L_0 be bounded at $y = 0$ by means of a perfectly conducting plane (Figure 4). We then have on this plane $E_z = ik\rho_0\,\partial\varphi/\partial y = 0$ and therefore the function φ is continued from the region $y > 0$ into the region $y < 0$ in an even fashion. In this connection. we arrive at an examination of the field in the upper half-plane outside the contour $L = L_0 + L_0^*$. where L_0^* is the mirror image of the contour L_0 about the z-axis. Taking into account the fact that the function φ is even in the case of complex amplitudes, we find

$$A(\pm w, u) = 2\int_{L_0} e^{-uz}\left[\left(\frac{\partial\varphi}{\partial n} + u\varphi\cos(n, z)\right)\cos wy + w\varphi\cos(n, y)\sin wy\right] dl. \tag{21}$$

We similarly reduce the integration along the contour L_0 in the general formula (15), to:

$$\varphi = \frac{i}{4}\int_{L_0}\left(G_0\frac{\partial\varphi}{\partial n} - \varphi\frac{\partial G_0}{\partial n}\right) dl,$$
$$G_0 = G(y, z, \eta, \zeta) + G(y, z, -\eta, \zeta). \tag{22}$$

2. Energy Characteristics

We have established earlier that the total field above the impedance plane and the surface-wave field are characterized by a single function $A(\pm w, u)$. Let us show now that we can express in terms of this function the energy characteristics. For this purpose we first calculate the average electromagnetic power carried by the surface waves through a unit transverse section at y → ±∞, starting from the general formula

$$P_{\pm} = \pm \frac{1}{2} \int_{0}^{\infty} \operatorname{Re} (E_z H_x^*)\, dz \quad \text{for} \quad y \to \pm \infty.$$

Substituting Equations (2) and (16) in this expression, we obtain the simple relation

$$P_{\pm} = \rho_0 \frac{pk^3}{4h} \left| A\,(\pm h, p) \right|^2. \tag{23}$$

On the other hand, in the case of unilateral radiation (Figure 4) we have for the total complex power

$$P_{\text{c}} = \int_{L_0} E_l H_x^*\, dl = i\rho_0 k^3 \int_{L_0} \varphi^* \frac{\partial \varphi}{\partial n}\, dl,$$

which can be represented in the form of a sum

$$P_{\text{к}} = P_{1\text{c}} + P_{2\text{c}} + P_{3\text{c}}, \quad P_{s\text{к}} = i\rho_0 k^3 \int_{L_0} \varphi_s^* \frac{\partial \varphi}{\partial n}\, dl \quad (s = 1, 2, 3). \tag{24}$$

For specified values of the function φ and its normal derivative on L_0, the individual terms of (24) are interpreted as before. Using furthermore the expressions in (19), we obtain for P_{2c} and P_{3c} the general formulas

$$\overset{*}{P}_{2\text{c}} = \rho_0 \frac{ik^3}{4\pi} \int_{-\infty}^{\infty} A\,(w, u)\, B\,(w, u) \frac{dw}{u}, \quad \overset{*}{P}_{3\text{c}} = \rho_0 \frac{ik^3 p}{2\pi} \int_{-\infty}^{\infty} \frac{A\,(w, u)\, B\,(w, u)}{u\,(u - p)}\, dw, \tag{25}$$

$$B\,(w, u) = \int_{L_0} \frac{\partial \varphi^*}{\partial n} e^{-iwy - uz}\, dl. \tag{26}$$

We now represent the complex powers in the customary form, putting $P_{sc} = 2P_{sa} + i2P_{sr}$ $(s = 1, 3)$, where P_{sa} and P_{sr} are respectively the active and reactive power; we then have for $P_{0a} = P_{1a} + P_{2a}$

$$P_{0a} = \rho_0 \frac{ik^3}{4} \int_{L_0} \left(\psi^* \frac{\partial \varphi}{\partial n} - \psi \frac{\partial \varphi^*}{\partial n} \right) dl, \quad \psi = \varphi_1 + \varphi_2.$$

The function ψ satisfies the conditions

$$\psi = \varphi, \quad \frac{\partial \psi}{\partial n} = \frac{\partial \varphi}{\partial n} \text{ on } L_0, \quad \left. \begin{array}{ll} \dfrac{\partial \psi}{\partial z} = 0 & \text{for } z = 0 \\ \dfrac{\partial \psi}{\partial y} = & \text{for } y = 0 \end{array} \right\} \text{outside of } L_0$$

Consequently, the preceding relation assumes the form

$$P_{0a} = \rho_0 \frac{ik^3}{4} \int_{L_0} \left(\psi^* \frac{\partial \psi}{\partial n} - \psi \frac{\partial \psi^*}{\partial n} \right) dl,$$

in which we can transform the contour L_0, on the basis of Green's formula, into a quarter-circle of large radius r, contiguous in part with the y- and z-axes. After making this transformation and taking (20) into account, we find

$$P_{0a} = \frac{\rho_0 k^3}{16\pi} \int_{0}^{\pi/2} \left| A\,(k \cos \theta, -ik \sin \theta) + A\,(k \cos \theta, ik \sin \theta) \right|^2 d\theta. \tag{27}$$

In the case of radiation in both directions, the contour L_0 in (24) and (26) must be replaced by the contour L. Accordingly, the upper limit of integration (27) is replaced by π.

II. APPROXIMATE METHODS OF CALCULATING THE SURFACE-WAVE FIELD

We now discuss, on the basis of the first part of this paper, approximate methods for the determination of the surface-wave field in various cases. The results obtained are illustrated with specific calculations.

1. General Considerations

Equations (15) and (17) enable us to indicate directly the general form of the approximate solution when the contour L is closed and the geometric center has a sufficient height z_0 above the impedance plane $z = 0$. In fact, let the function φ_∞ be the exact solution in free space; we can then write as the first approximation

$$\varphi \simeq \frac{i}{4} \int\limits_L \left(G \frac{\partial \varphi_\infty}{\partial n} - \varphi_\infty \frac{\partial G}{\partial n} \right) dl, \tag{28}$$

$$A(\pm w, u) \simeq e^{-uz_0} \int\limits_L e^{\pm iwy - uz} \left[\frac{\partial \varphi_\infty}{\partial n} - \varphi_\infty (\pm iw \cos(n, y) - u \cos(n, z)) \right] dl \tag{29}$$

$$(w = k \operatorname{ch} \alpha,\ u = k \operatorname{sh} \alpha),$$

where the ordinate z in (29) is measured from the height of the geometric center of the contour L. It is obvious that this approximation is the more accurate, the higher the cylindrical surface above the impedance plane $z = 0$. Thus, exact solutions for free space can be used directly for approximate calculations of the wave field in the case under consideration. We can, however, establish a closer connection between φ_∞ and the approximate solution. Indeed, according to (20), we have for φ_∞ the asymptotic expression

$$\varphi_\infty = \frac{iQ(k, \theta)}{(8\pi kr)^{1/2}} e^{-i\left(kr - \frac{\pi}{4}\right)}, \quad Q(k, \theta) = \int\limits_L \left(f \frac{\partial \varphi_\infty}{\partial n} - \varphi_\infty \frac{\partial f}{\partial n} \right) dl, \tag{30}$$

$$f = e^{ik(y \cos\theta + z \sin\theta)},$$

where the asymptotic characteristic $Q(k, \theta)$ determines the directivity pattern of the radiated or scattered waves in free space. Comparing (29) with (30), we obtain

$$A(w, u) = e^{-uz_0} Q(k, i\alpha),\ A(-w, u) = e^{-uz_0} Q(k, \pi - i\alpha). \tag{31}$$

In particular, for $\alpha = \alpha_0$, when $w = h$ and $u = p$, we obtain from (31) the complex amplitudes of the surface waves. Consequently, for an approximate determination of the complex amplitude $A(\pm w, u)$ it is sufficient to specify only one asymptotic characteristic $Q(k, \theta)$ of the radiated or scattered cylindrical waves in free space.

To illustrate the general considerations, let us consider several particular cases. Let for example, the electromagnetic field be excited above the impedance plane $z = 0$ by a semi-infinite flat waveguide with perfectly conducting walls (Figure 5). We then have for $Q(k, \theta)$ the expression

$$Q(k, \theta) = \int\limits_{-a}^{a} \left(\frac{\partial \varphi_\infty}{\partial y} - ik \cos\theta \varphi_\infty \right) e^{ikz \sin\theta} dz -$$

$$- ik \sin\theta \int\limits_{-\infty}^{0} e^{iky\cos\theta} \left(\varphi_\infty(y, a)\, e^{ika\sin\theta} - \varphi_\infty(y, -a)\, e^{-ika\sin\theta} \right) dy. \tag{32}$$

The first term in (32) is connected with the structure of the field on the open end of the waveguide, and the second term is due to the electric currents that flow on the outer walls of the waveguide. The exact values of $Q(k, \theta)$ for different types of E waves are calculated in [6], where it is shown, in particular, that for certain types of E waves we can use the Kirchhoff method to calculate $Q(k, \theta)$, i.e., to neglect currents flowing on the outer walls of the waveguide, and to consider the structure of the field at the open end to be the same as in an infinite flat waveguide. Using the Kirchhof approximation, we can put $\partial\varphi/\partial y = -ih_s\varphi$ with $y = 0$; we then obtain from (31) and (32)

$$A(\pm h, p) = \frac{1}{ik\rho_0} e^{-pz_0} \left(1 \pm \frac{h}{h_s} \right) \int\limits_{-a}^{a} E_s(z)\, e^{-pz} dz, \tag{33}$$

where h_s is the propagation constant in the waveguide; $E_s(z)$ is the value of the E_z component of the electric field intensity for the given E wave in the waveguide.

It is obvious that the Kirchhoff approximation can be used in the most general case for arbitrary form of the contour L, assuming that $\partial\varphi_\infty/\partial n = -ih_s\varphi_\infty$ on the slots and neglecting the currents flowing on the perfectly conducting parts of this contour. Here, as usual, the approximation is more accurate the wider the slots are compared with the wavelength in free space. In all cases, when using the exact or approximate solution for φ_∞, it is essential to have the distortions of the field near the contour L, due to the impedance plane, quite weak, something which is possible when z_0, the height of the geometric contour L_0, is sufficiently large.

We now consider a second example. We assume that the surface waves are produced on the impedance plane as a result of diffraction of normally incident plane waves by a perfectly conducting cylinder (H-polarization). We separate from the total field the waves incident on and reflected from the impedance plane.

$$\varphi_0 = e^{ikz_0\sin\beta}\varphi_{01} + Re^{-ikz_0\sin\beta}\varphi_{02}, \quad R = \frac{ik\sin\beta + p}{ik\sin\beta - p}, \qquad (34)$$
$$\varphi_{01} = e^{-ik(y\cos\beta - z\sin\beta)}, \quad \varphi_{02} = e^{-ik(y\cos\beta + z\sin\beta)}$$

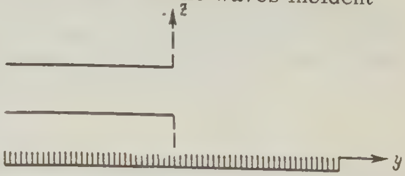


Figure 5.

where β is the elevation angle; R is the coefficient of reflection, and the ordinate z, as before, is reckoned from the height z_0 of the geometric center of the contour L.

It is obvious that the function φ_∞ can be represented as a sum

$$\varphi_\infty = e^{ikz_0\sin\beta}\varphi_1 + Re^{-ikz_0\sin\beta}\varphi_2, \qquad (35)$$

where the functions φ_1 and φ_2 satisfy the boundary conditions

$$\frac{\partial\varphi_1}{\partial n} = -\frac{\partial\varphi_{01}}{\partial n}, \quad \frac{\partial\varphi_2}{\partial n} = -\frac{\partial\varphi_{02}}{\partial n} \text{ at } L. \qquad (36)$$

In accordance with (30), (31), and (35), we find

$$\begin{aligned} A(w, u) &= e^{(ik\sin\beta - u)z_0}Q_1(k, i\alpha) + Re^{-(ik\sin\beta + u)z_0}Q_2(k, i\alpha), \\ A(-w, u) &= e^{(ik\sin\beta - u)z_0}Q_1(k, \pi - i\alpha) + Re^{-(ik\sin\beta + u)z_0}Q_2(k, \pi - i\alpha), \\ Q_{1,2}(k, \theta) &= \int_L \left(f\frac{\partial\varphi_{1,2}}{\partial n} - \varphi_{1,2}\frac{\partial f}{\partial n}\right)dl. \end{aligned} \qquad (37)$$

The expressions in (37) become particularly simple in the case of a circular cylinder. In this case $Q_1(k, \theta) = Q_0(k, \theta + \beta)$ and $Q_2(k, \theta) = Q_0(k, \theta - \beta)$, where $Q_0(k, \theta)$ is the asymptotic characteristic of the waves scattered from the circular cylinder as a result of diffraction of plane waves incident along the polar axis. Further, if we put $\beta = i\alpha_0$, i.e., when $k\cosh\alpha_0 = h$, $k\sinh\alpha_0 = p$, and $R = 0$, then the incident wave turns into a surface wave, and we have now for the circular cylinder

$$\begin{aligned} A(w, u) &= e^{-(u+p)z_0}Q_0(k, i(\alpha, + \alpha_0)), \quad A(-w, u) = \\ &= e^{-(u+p)z_0}Q_0(k, \pi + i(\alpha_0 - \alpha)). \end{aligned} \qquad (38)$$

In particular we obtain for the complex amplitudes $A(\pm h, p)$ of the surface waves ($\alpha = \alpha_0$) the following interesting result:

$$A(h, p) = e^{-2pz_0}Q_0(k, 2i\alpha_0), \; A(-h, p) = e^{-2pz_0}Q_0(k, \pi). \qquad (39)$$

We see, therefore, that the complex amplitude $A(-h, p)$ of the reflected surface waves coincides, apart from the factor $\exp(-2pz_0)$, with the asymptotic characteristic $Q_0(k, \pi)$ of the backward scattering under normal incidents of H-polarized plane homogeneous waves on a circular cylinder in free space.

2. Slot and Waveguide Excitation of Surface Waves

Let us consider several examples, for which the wave field can be determined in a different fashion. Assume that we have a slot in a perfectly conducting vertical plane, bearing against the impedance plane $z = 0$ (Figure 6). In this case $E_z = ik\rho_0\partial\varphi/\partial y$ in the

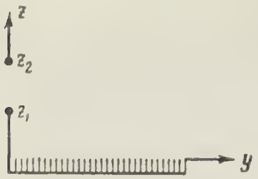


Figure 6

gap (z_1, z_2) on the z-axis at $y = +0$, where z_1 and z_2 are the ordinates of the edges of the slot. From the general formulas (21), (22), and (26) we obtain directly

$$\varphi = \frac{1}{2k\rho_0}\int_{z_1}^{z_2} E_z(\zeta)\, G(y, z, 0, \zeta)\, d\zeta, \tag{40}$$

$$A(\pm w, u) = \frac{2}{ik\rho_0}\int_{z_1}^{z_2} E_z(z)\, e^{-uz}\, dz,$$

$$B(w, u) = -\frac{1}{ik\rho_0}\int_{z_1}^{z_2} E_z^*(z)\, e^{-uz}\, dz. \tag{41}$$

It is clear, therefore, that the excited electromagnetic field and its energy characteristics can be completely determined if the function $H_Z(z)$ is specified. In particular, for a concentrated slot $E_Z = V_0\delta(z - z_0)$ and the foregoing formulas simplify to

$$\varphi = \frac{V_0}{2k\rho_0} G(y, z, 0, z_0), \quad A(\pm w, u) = \frac{2V_0}{ik\rho_0} e^{-uz_0}, \quad B(w, u) = -\frac{V_0^*}{ik\rho_0} e^{-uz_0}. \tag{42}$$

Let us calculate the energy characteristics in this very simple case. First of all, by virtue of (23) we obtain for the average power and radiation conductance of the surface waves

$$P_\pm = \frac{kp}{\rho_0 h} |V_0|^2 e^{-2pz_0}, \quad g_0 = \frac{2P_\pm}{|V_0|^2} = \frac{2kp}{\rho_0 h} e^{-2pz_0}. \tag{43}$$

we then obtain on the basis of (10) and (25)

$$g_{2\kappa} = \frac{P_{2c}}{|V_0|^2} = \frac{k}{2\rho_0} H_0^{(1)}(\nu) \quad (\nu = 2kz_0), \tag{44}$$

where g_{2c} is the additional complex admittance, which appears in the presence of a perfectly conducting plane $z = 0$. In precisely the same manner we readily obtain from (27)

$$g_{0a} = \frac{2P_{0a}}{|V_0|^2} = \frac{k}{2\rho_0}(1 + J_0(\nu)). \tag{45}$$

Here g_0 is the active radiation conductance of the concentrated slot above a perfectly conducting plane, the first term corresponding to the active conductance in free space, and the second part of the active conductance, from (42) in the second formula of (25), we arrive at an expression for the third part of the complex admittance

$$g_{3c}^* = \frac{P_{3c}^*}{|V_0|^2} = \frac{2ikp}{\rho_0\pi}\int_0^\infty \frac{e^{-2uz_0}}{u(u-p)}\, dw \quad (u = (w^2 - k^2)^{1/2}).$$

Let us consider the integration paths $(0, k)$ and (k, ∞) separately. On the first path we have $u = i(k^2 - w^2)^{1/2}$, and on the second the point $w = h$ is a pole which is to be circuited from above. Therefore, putting on the first part $w = k \sin\theta$ and going around the aforementioned pole, we can represent the expression for g_{3c} in the form

$$g_{3c} = g_0 - \frac{2k}{\rho_0\pi} Z \int_0^{\pi/2} \frac{e^{i\nu\cos\theta}}{Z + i\cos\theta}\, d\theta - \frac{2ikp}{\rho_0\pi}\int_0^\infty \frac{e^{-2uz_0}}{u-p}\,\frac{du}{w} \tag{46}$$

$$(Z = p/k, \quad w = (k^2 + u^2)^{1/2}),$$

where g_0 is the radiation conductance of the surface waves, defined by (43). We consider the principal value of the second integral in (46) in the Cauchy sense. It is easily seen that the active part of the admittance $g_{3a} = \operatorname{Re} g_{3c}$ is given by the relation

$$g_{3a} = g_0 - \frac{kZ}{2\rho_0\pi}\int_0^{2\pi} \frac{e^{i\nu\cos\theta}}{Z + i\cos\theta}\, d\theta.$$

Putting now $i \exp i\theta = t$, we arrive at a contour integral along the unit circle in the complex plane t, taken counterclockwise. Considering further the well-known equality (l is a closed contour surrounding the point $t = 0$)

$$\int_l t^{n-1} \exp \frac{\nu}{2}\left(t - \frac{1}{t}\right) dt = 2\pi i (-1)^n J_n(\nu),$$

we obtain the expression

$$g_{3a} = g_0 - \frac{kp}{\rho_0 h}\left(J_0(\nu) + 2\sum_{n=1}^{\infty} t_1^n J_n(\nu)\right) \quad (t_1 = (Z^2 + 1)^{1/2} - Z). \tag{47}$$

We thus have from (45) and (47) the following expression for the total active conductance

$$g_a = \frac{k}{2\rho_0}\left[1 + J_0(\nu) + \frac{4p}{h} e^{-2pz_0} - \frac{2p}{h}\left(J_0(\nu) + 2\sum_{n=1}^{\infty} t_1^n J_n(\nu)\right)\right]. \tag{48}$$

Formula (48) is suitable for calculations with moderate values of $\nu = 2kz_0$. In particular, when $\nu \ll 1$ it is seen from (48) that $g_0/g_a \simeq 2ph^{-1}(1 + ph^{-1})$, and in this case $g_a \simeq$ $\simeq k\rho_0^{-1}(1 + ph^{-1})^{-1}$. When $\nu \gg 1$, the excitation of surface waves is weak. A corresponding estimate can be obtained by the stationary-phase method, according to which

$$\int_0^{\pi/2} \frac{e^{i\nu\cos\theta}}{Z + i\cos\theta} d\theta = \left(\frac{\pi}{2\nu}\right)^{1/2} \frac{e^{i\left(\nu - \frac{\pi}{4}\right)}}{Z + i},$$

and therefore when $\nu \gg 1$ we have from (45) and (46)

$$\begin{aligned} g_a = g_0 + \frac{k}{2\rho_0}\left(1 + \left(\frac{2}{\pi\nu}\right)^{1/2} \cos\left(\nu - \frac{\pi}{4}\right)\right) - \\ - \left(\frac{2}{\pi\nu}\right)^{1/2} \frac{kZ}{\rho_0(Z^2 + 1)}\left(Z\cos\left(\nu - \frac{\pi}{4}\right) + \sin\left(\nu - \frac{\pi}{4}\right)\right). \end{aligned} \tag{49}$$

As can be seen from (40) and (41), a change to relations (42) for the concentrated slot is possible when the conditions $kb \ll 1$ and $pb \ll 1$ are simultaneously satisfied with $b =$ $= z_2 - z_1$ the width of the slot. An account of the finite dimensions of the slot calls for a detailed analysis of the field for $y < 0$, and subsequent matching of the results in the gap (z_1, z_2) of the z-axis. Let us assume that the gap is the open end of a flat waveguide, the lower boundary of which is an impedance plane with characteristic value $p_1 =$ $= kZ_1 > 0$, while the upper boundary is a perfectly conducting plane (Figure 7). The functions $\varphi_\pm(y, z)$, which determine the fields for $y > 0$ and $y < 0$, satisfy the continuity conditions of the points of the slot

$$\frac{\partial \varphi_-}{\partial y} = \frac{\partial \varphi_+}{\partial y}, \quad \varphi_- = \varphi_+. \tag{50}$$

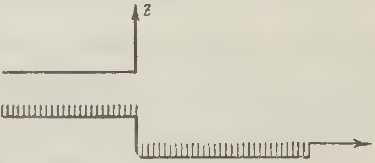


Figure 7

Consequently the characteristics of the field can be determined for $y > 0$ from the formulas (40) and (41) have the form

$$\varphi_+ = \frac{i}{2}\int_{z_1}^{z_2} \frac{\partial \varphi_-}{\partial y} G(y, z, 0, \zeta)\, d\zeta, \tag{51}$$

$$\begin{aligned} A(\pm w, u) &= 2\int_{z_1}^{z_2} \frac{\partial \varphi_-}{\partial y} e^{-uz}\, dz, \\ B(w, u) &= \int_{z_1}^{z_2} \frac{\partial \overset{*}{\varphi}_-}{\partial y} e^{-uz}\, dz. \end{aligned} \tag{52}$$

Inside the waveguide, for $y < 0$, the field is characterized by the system of waves

$$\varphi_n(y, z) = \psi_n(z) e^{\pm i h_n y}, \quad h_n = (k^2 - q_n^2)^{1/2}, \quad \psi_n(z) = \cos q_n(z - z_2), \tag{53}$$

where q_n are the roots of the equation $q_n \tan q_n b = -p_1 (2n - 1)\ \pi/2 < q_n b < n\pi\ n = 1, 2, \ldots,$ where $q_0 = ip_0$; p_0 is the positive root of the equation $p_0 \tanh p_0 b = p_1$, corresponding to a slow wave with propagation constant h_0. We note also that when $p_1 = 0$, i.e., when the boundary $z = z_1$ is a perfectly conducting plane, we have the usual flat waveguide and here already $p_0 = 0 (h_0 = k)$ and all $q_n b = n\pi (n = 1, 2, \dots)$.

Consequently, in the rigorous formulation, the field in the waveguide is specified in the form

$$\varphi_- = \psi_s(z) e^{-ih_s y} + \sum_{n=0}^{\infty} a_n \psi_n(z) e^{ih_n y}, \tag{54}$$

where the first term determines the incident wave ($k > q_s$) and the second term is the sum of the reflected waves together with the exponentially damped part of the field. From (54) we obtain for $y = 0$

$$\varphi_- = \psi_s(z) + \sum_{n=0}^{\infty} a_n \psi_n(z), \quad \frac{\partial \varphi_-}{\partial y} = -ih_s \psi_s(z) + i \sum_{n=0}^{\infty} h_n a_n \psi_n(z). \tag{55}$$

Substituting the second of these expressions in (51) and then using the boundary condition (50) and also the orthogonality of the functions $\psi_n(z)$, we obtain an infinite system of equations for the determination of the coefficients a_n.

This method of setting up the exact solution is cumbersome, since, the use of the second condition in (50) entails cumbersome calculations of the coefficients of the expansion in the functions $\psi_n(z)$ and, it is necessary to determine whether the infinite system of equations for the coefficients a_n can be solved. Taking these features into account, we shall point out only a few limiting cases. First of all, for a wide slot ($kb \gg 1$), we can neglect, within the framework of the Kirchhoff approximation, the effect of reflection and consequently put $\partial\varphi_-/\partial y \simeq -ih_s\psi_s(z)$ for $y = 0$. Substituting this expression in (52) we are in a position to calculate approximately the complex amplitude of the excited surface waves, together with the energy characteristics of the total electromagnetic field. Obviously, the second condition of (50) cannot be satisfied in strict form in this approximation. Therefore for this approximation to be accuract it is necessary not only to satisfy the condition $kb \gg 1$, but also to have the center of the slot $z_0 = (z_1 + z_2)/2$, high enough above the impedance plane, so that the distortion of the field at the points of the slot, due to this plane, is small.

Now let kb be small, more accurately $kb < \pi/2$ when $p_1 \neq 0$ or $kb < \pi$ when $p_1 = 0$ In this case all h_n ($n = 1, 2, \dots$) are pure imaginary, and the incident and reflected waves correspond to $s = 0$. Neglecting in (54) the exponentially damped part of the field, we obtain for $y = 0$ the approximate expressions

$$\varphi_- \simeq (1 + a_0) \psi_0(z), \quad \frac{\partial \varphi_-}{\partial y} \simeq -ih_0 (1 - a_0) \psi_0(z). \tag{56}$$

Substituting the second expression of (56) in (51) and (52) we have

$$\varphi_+(y, z) \simeq \frac{h_0}{2} (1 - a_0) \int_{z_1}^{z_2} \psi_0(\zeta) G(y, z, 0, \zeta) d\zeta, \tag{57}$$

$$A(\pm w, u) \simeq -2ih_0 (1 - a_0) \int_{z_1}^{z_2} \psi_0(z) e^{-uz} dz,$$

$$B(w, u) \simeq ih_0 (1 - a_0^*) \int_{z_1}^{z_2} \psi_0(z) e^{-uz} dz, \tag{58}$$

where the second condition of (50) can be satisfied only in the mean, so that the coefficient of reflection a_0 is only approximate. This coefficient can also be determined from energy considerations or else by stipulating for a narrow slit ($kb \ll 1$) the satisfaction of the second condition of (50) at the point $z + z_0$.

To highlight the difference between the approximations (57) or (58) and the results for a concentrated slot, let us calculate the radiation conductances of the surface waves in the simplest case $p_1 = 0$, i.e., when $\psi_0(z) = 1$ and $h_0 = k$. From (58) we obtain

$$A_-(\pm h, p) = \frac{2V_0}{ik\rho_0} e^{-pz_0} \frac{\sinh(pb/2)}{pb/2}, \quad V_0 = \int_{z_1}^{z_2} E_z dz = \rho_0 k^2 (1 - a_0) b, \tag{59}$$

where V_0 can be regarded as a voltage applied to the edges of the slot. We now use (23) and obtain the following expression for the radiation conductance of the surface waves

$$g_0 = \frac{2P_\pm}{|V_0|^2} = \frac{2kp}{\rho_0 h} e^{-2pz_0} \left(\frac{\operatorname{sh}(pb/2)}{pb/2} \right)^2. \tag{60}$$

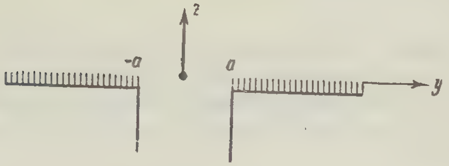


Figure 8

A comparison of (43) with (60) again shows that to change over to a concentrated slot it is necessary to satisfy the condition $pb \ll 1$. At the same time, the approximate expression (60) obtained for $kb < \pi$ ($p_1 = 0$), indicates that the conductance is greater than that for a concentrated slot.

Now let the sources be distributed on the segment (-a, a) of the y-axis; then, shrinking the contour L in (17) to this segment, we have

$$A(\pm h, p) = \int_{-a}^{a} \left(\frac{\partial \varphi}{\partial z} + p\varphi \right) e^{\pm ihy} dy. \tag{61}$$

If the segment (-a, a) is the open end of a flat waveguide (Figure 8) then for $kb \gg 1$ we can proceed as previously, namely put for $y = 0$

$$\frac{\partial \varphi}{\partial z} + p\varphi = -\frac{1}{ik\rho_0} \left(1 + \frac{p_i}{h_s} \right) E_s, \quad E_s = E_0 \cos q_s (y - a) \tag{62}$$

and thus calculate approximately the surface-wave field. It is obvious that this approximation is less accurate than the one considered above, since a noticeable field distortion may take place in the vicinity of the points $y = \pm a$, owing to the presence of the impedance plane.

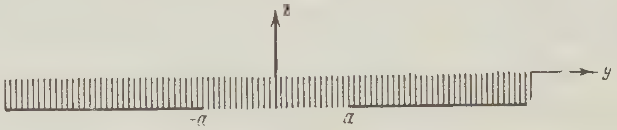


Figure 9.

Another variant of the excitation of surface waves is indicated in Figure 9. Here the impedance plane is realized with the aid of a comb structure with an open section (-a, a), to which the specified field $E_y = E_S(y)$ is applied with $z = -l$. Assuming, as is customary, that when $-l < z < 0$ the field in the grooves is characterized by a function $\varphi_i = C_1 \exp ikz + C_2 \exp(-ikz)$, we can readily establish that

$$\frac{\partial \varphi}{\partial z} + p\varphi = \frac{i}{k\rho_0} \sec kl\, E_s \quad \text{for} \quad z = 0, \ |y| < a.$$

Obviously, we have expressions of the same type for the surface-wave amplitudes in either excitation variant. On the other hand, if the segment (-a, a) is perfectly conducting, and has concentrated slots at the points y_s ($s = 0, 1, \ldots, m$) (Figure 10, (61) assumes the form

$$A(\pm h, p) = -\frac{1}{ik\rho_0} \sum_{s=0}^{m} V_s e^{\pm ihy_s} + p \int_{-a}^{a} \varphi e^{\pm ihy} dy, \tag{63}$$

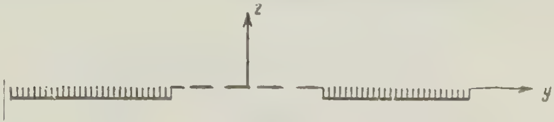


Figure 10

in which the values of the function φ on the segment (-a, a) are not known exactly, but are given in first approximation by

$$\varphi \simeq \varphi_0 = -\frac{1}{2k\rho_0}\sum_{s=0}^{m} V_s H_0^{(2)}\left(k\left((y-y_s)^2+z^2\right)^{1/2}\right). \tag{64}$$

This approximation, as it were, "takes no notice" of the impedance plane that begins at $y = \pm a$, and can be used in the zone where the slot is situated, provided the outermost slots are relatively far from the points $y = \pm a$. For the sake of illustration we consider a single concentrated slot at the point $y_0 = 0$. We then obtain from (63) and (64) ($\alpha = ha$, $\beta_0 = k/h$)

$$A(\pm h, p) = -\frac{V_0}{ik\rho_0}\left(1 + i(1-\beta_0^2)^{1/2} H_c(\beta_0, \alpha)\right),$$

$$H_c(\beta_0, \alpha) = \int_0^{\alpha} \cos t H_0^{(2)}(\beta_0, t)\, dt. \tag{65}$$

Here β_0 is the retardation coefficient, and function $H_c(\beta_0,\alpha)$ is tabulated in reference [7]. On the basis of (23), we obtain an expression for the two-sided radiation conductance of the surface waves

$$g_0 = \frac{k}{\rho_0}(1-\beta_0^2)^{1/2}\left|1 + i(1-\beta_0^2)^{1/2} H_c(\beta_0, \alpha)\right|^2. \tag{66}$$

Here $H_c(\beta_0, \infty) = i(1-\beta_0^2)^{-1/2}$ and therefore $g_0 = 0$ when $\alpha = \infty$, as it should be.

For a more rigorous analysis of the edge effect, we point to the exact solution of this problem for the particular case when the slot is located at the point $y = -y_0$ of a perfectly conducting semiinfinite portion $(-\infty, 0)$. of the y-axis, where the second semiinfinite portion $(0, \infty)$, is the impedance plane. We represent the function φ in the form of a sum

$$\varphi = \varphi_0 + \psi, \quad \varphi_0 = -\frac{V_0}{2k\rho_0} H_0^{(2)}\left(k\left((y+y_0)^2+z^2\right)^{1/2}\right). \tag{67}$$

We then obtain on the basis of (3) ($y< 0<$ and the condition $\partial\varphi/\partial z = 0$ for $z = 0$ and $y < 0$ ($y \neq -y_0$) the relations

$$\frac{\partial\psi}{\partial z} = 0 \quad \text{for} \quad z=0,\ y<0, \quad \frac{\partial\psi}{\partial z} + p\psi = \frac{pV_0}{2k\rho_0} H_0^{(2)}(k(y+y_0)) \quad \text{for} \quad z=0,\ y>0.$$

Substituting the plane-wave expansion of the function $\psi(y, z)$

$$\psi(y, z) = \int_{-\infty}^{\infty} \Gamma(w)\, e^{-iwy-uz}\frac{dw}{u} \qquad (u = (w^2-k^2)^{1/2},\ \operatorname{Re} u > 0) \tag{68}$$

and using the integral representation of the function $H_0^{(2)}(x)$, we obtain

$$\int_{-\infty}^{\infty} \Gamma(w)\, e^{-iwy}\, dw = 0 \quad \text{for} \quad y < 0, \tag{69}$$

$$\int_{-\infty}^{\infty}\left(\frac{p}{u}-1\right)\Gamma(w)\, e^{-iwy}\, dw = -\frac{pV_0}{\pi i k\rho_0}\int_0^{\infty} e^{-ik(y+y_0)\operatorname{ch} t}\, dt \quad \text{for} \quad y > 0\ (\operatorname{Im} k<0). \tag{70}$$

The theory of equations of this type [8 — 12] is based on factoring the expression $p/u - 1 = f_+(w)\, f_-(w)$, where the function $f_+(w)$ is regular and has no zeros in the upper half plane ($\operatorname{Im} w \geq 0$), the function $f_-(w)$ is regular and has no zeros in the lower half plane ($\operatorname{Im} w \leq 0$). These functions are known [12], and for the simplest right-hand part, $v = r \exp(-isy)$ ($\operatorname{Im} s < 0$), the solution is given by

$$\Gamma_s(w) = -\frac{r}{2\pi i}\frac{1}{f_-(s)\, f_+(w)\, (w-s)},$$

while summation must be carried out for the right half in (70) ; as a result we get

$$\Gamma(w) = -\frac{pV_0}{2\pi^2 k\rho_0 f_+(w)} \int_0^\infty \frac{e^{-ky_0 \operatorname{ch} t}}{f_-(k \operatorname{ch} t)(w - k \operatorname{ch} t)}. \tag{71}$$

Formulas (67) , (68) , and (71) give an exact solution of this problem, and the separation of the residues at the point $f_+(w)$ determines the excited surface-wave field for $y \to \infty$.

REFERENCES

1. M.D. Khaskind, Propagation of Sound and Electromagnetic Waves in a Half Space, Akusticheskiy zhurnal (Acoustic Journal), 1959, 5, 4, 464.
2. G.D. Malyuzhinets, Concerning One Generalization of the Weyl Formula for the Wave Field above an Absorbing Plane, Doklady AN SSSR (New Series), 1948, 60, 3, 367.
3. M.D. Khaskind, Excitation of Surface Electromagnetic Waves on Plane Dielectric Coatings, Radiotekhnika i elektronika, 1960, 5, 2, 188.
4. M.D. Khaskind, Approximate Method for Determining Hydrodynamic Pumping Characteristics, Izv. AN SSSR, Division of Technical Sciences, 1954, 11, 66.
5. M.D. Khaskind, Perturbing Forces and Floodability of Ships in Sea Waves, ibid 1957, 7, 65.
6. L.A. Vaynshteyn, Difraktsiya elektromagnitnykh i zvukovykh voln na otkrytom kontse volnovoda, Izd. Sovetskoye radio, 1953. (Diffraction of Electromagnetic and Sound Waves by an Open End of a Waveguide), Soviet Radio Press, 1953.
7. L. Schwartz, Untersuchung einiger mit Zylinderfunctionen nullter Ordnung Verwandter Funktionen, Luftfahrtforschung, 1943, 20.
8. N. Wiener, E. Hopf, Über eine Klasse singularer Integralgleichungen, Sitz, Berlin, Akad. Wiss., 1931, 696-706.
9. E.C. Titchmarsh, Introduction to the Theory of Fourier Integrals, Clarendon Press, Oxford, 1937.
10. V.A. Fok, O nekotorykh integral'nykh uravneniyakh matematicheskoy fiziki, (Concerning Certain Integral Equations of Mathematical Physics), Matem. sbornik, 1944, 14, 1.
11. V. I. Smirnov, Kurs vysshey matematiki, (Course of Higher Mathematics), v. 4, GTI, 1951, pp. 179 — 188.
12. G.A. Grinberg, V.A. Fok, On the Theory of Shore Refraction of Electromagnetic Waves. Issledovaniya po rasprostraneniyu radiovoln, (Research on Radio Wave Propagation), Collection no. 2, Izd. AN SSSR, 1948, 69.

Submitted to the editors 16 November 1960.

DIFFRACTION OF ELECTROMAGNETIC WAVES BY CONCAVE SURFACE OF CIRCULAR CYLINDER

B. Ye. Kinber

The diffraction of the Sommerfeld edge wave on the concave surface of a mirror leads to multiple reflections of this wave and to the appearance of "whispering gallery" waves. These effects are investigated in the present article using a model consisting of a current filament placed inside the edge cylinder. The radiation conditions are satisfied by introducing multiple sheets. The laws of geometric optics and the field near the caustic of the whispering-gallery wave are investigated.

INTRODUCTION

The presently known solutions of diffraction problems pertain in the overwhelming majority of cases to convex bodies [1, 2]. In the limit as $\lambda \to 0$, they go into the solutions of geometrical optics, which yields, under these conditions, singly reflected and divergent beams of rays. It is known [3, 4] that diffraction by concave bodies gives rise to qualitative new effects such as caustic field surfaces, multiple reflections, the whispering-gallery effect, etc.

The investigated cases of diffraction on concave bodies — parabolic cylinder [5] and paraboloid of revolution [6] with the source of excitation located at the focus — are not typical cases, for in the limit they yield a single reflection. The singularities of diffraction by convex bodies manifest themselves when the source is located near the wall and the geometric-optic solution yields multiple reflections.

The analysis of the diffraction effect is made complicated by the fact that, owing to technical difficulties, the calculations can be carried out only for bodies of simplest forms, in which only closed cavities have concave surfaces (internal surface of a cylinder, sphere, ellipsoid, etc.).

The excitation of a closed volume by a source placed inside the volume produces standing waves, rather than the traveling waves that are characteristic of the exterior problem of diffraction theory.

In reference [2], waves that circulate over the surface of the cylinder many times are considered formally as being located not on an ordinary surface, but on a Riemann surface. From this point of view, the wave excited by the source circles the origin and never returns to the point of excitation.

If this procedure is adopted for internal problems, it can be assumed that the field excited by a source located inside a closed volume satisfies the radiation principle. In this case, however, the diffraction field of interest to us should be considered on each sheet of the surface separately.

In the present paper we consider the simples problem of diffraction on a concave surface — the internal surface of a circular cylinder. The source of the field is a current filament. This problem can be regarded also as the limiting case of an infinitely thin helical waveguide of rectangular cross section with an internal wall that has zero radius of curvature and with a helix that has a zero pitch. To solve this problem, we introduce a system of normal waves traveling in the direction $\pm\varphi$. This system of waves is equivalent to a system of guided waves for helical waveguides. We then construct the Green's function of the problem. An analysis of the resultant solution enables us to describe the multiple reflections of the edge wave, the field near the caustics, and the whispering-gallery effect on the concave surface of a cylindrical mirror of finite size.

1. EIGENFUNCTIONS OF THE PROBLEM

Let us determine the eigenfunctions of the wave equation

$$\Delta U + k^2 U = 0, \tag{1}$$

where U is the Hertz potential* for the internal region of a circular cylinder of radius a, located on a plane having an infinite number of sheets, subject to the usual boundary conditions

$$U = 0, \tag{2}$$

$$\frac{\partial U}{\partial r} = 0. \tag{3}$$

Introducing the cylindrical coordinates r and φ and separating variables, we obtain

$$\Phi'' + \nu^2 \Phi = 0, \tag{4}$$

$$(kr)^2 Z'' + (kr) Z' + [(kr)^2 - \nu^2] Z = 0, \tag{5}$$

$$U = \Phi(\varphi) Z(kr), \tag{6}$$

where ν is the separation constant.

The eigenfunctions and the eigenvalues ν_s are determined by solving (5) subject to boundary conditions (2) or (3).

Let us consider first the solution of (5), $Z_\nu = \alpha J_\nu(kr) + \beta J_{-\nu}(kr)$, (for the sake of uniqueness, we put Re $\nu > 0$) for the field between two coaxial cylinders of radii a and b, and then go to the limit as $b \to 0$. This leads to an equation

$$J_\nu(ka) J_{-\nu}(kb) - J_{-\nu}(ka) J_\nu(kb) = 0$$

subject to boundary condition (2) and to an equation

$$J'_\nu(ka) J'_{-\nu}(kb) - J'_{-\nu}(ka) J'_\nu(kb) = 0$$

with boundary condition (3).

Going to the limit as $b \to 0$ and recognizing that $J_m(x) \sim x^m$ when $x \ll 1$, we find that the eigenvalues ν_s are determined respectively as the roots of the equations

$$J_\nu(ka) = 0, \tag{7}$$

$$J'_\nu(ka) = 0 \tag{8}$$

subject to boundary conditions (2) and (3) and to the additional condition Re $\nu > 0$,

Equation (7) or (8) has for Re $\nu > 0$ an infinite number of roots, of which approximately $2a/\lambda$ are real ($0 \leqslant \nu \leqslant ka$) and correspond to undamped guided waves. The remaining roots are imaginary and correspond to damped waves.

The two independent solutions of (4)

$$\Phi_1 = e^{i\nu\varphi},$$

$$\Phi_2 = e^{-i\nu\varphi} \tag{9}$$

correspond to two systems of waves, traveling over the angle φ in opposite directions.

The condition Im $\nu > 0$ follows from the radiation principle. Consequently, all the eigenvalues of (7) or (8) are contained in the first quadrant of the complex plane ν (Figure 1).

The field of each natural wave is "pressed" against the concave wall of the cylinder (the Krasnushkin effect [4]). When $kr < \nu$ the field decreases monotonically with increasing distance from the wall, i.e., with decreasing r (Figure 2).

The structure of the natural waves of the concave surface is close to the structure of the natural waves for a flat-layered medium in the case of the earth's surface waveguide [8] /

In the region $kr > \nu$ the quantity ν/r is the component in the φ direction, and the quantity

---

*Since the field vectors are proportional to U or its derivatives, we shall henceforth consider the function U itself.

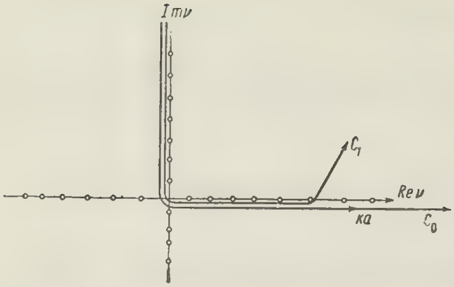


Figure 1. Location of roots of the equation $J_\nu(ka) = 0$ [or $J'_\nu(ka) = 0$] on the complex ν-plane. The eigenvalues of the problem correspond to the roots located in the first quadrant. The largest real root is $\simeq ka$.

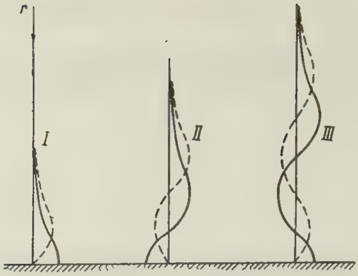


Figure 2. Profiles of three first normal waves. The lower line corresponds to the cylinder wall r = a. The first proper wave corresponds to the largest real ν. The continuous line is the boundary condition $dU/dr = 0$, the dotted line is the boundary condition $U = 0$.

$\sqrt{k^2 - \left(\frac{\nu}{r}\right)^2}$ is the r-component of the wave vector k. This follows from the asymptotic expansion for Bessel functions

$$J_\nu(kr) = \sqrt{\frac{2}{\pi}} \frac{\cos \xi}{\sqrt[4]{(kr)^2 - \nu^2}}, \quad |kr - \nu| > (kr)^{1/3}, \tag{10}$$

where

$$\xi = \sqrt{(kr)^2 - \nu^2} - \nu \arccos \frac{\nu}{kr} - \frac{\pi}{4}, \tag{11}$$

hence

$$k_r = \frac{d\xi}{dr} = \sqrt{(k)^2 - \left(\frac{\nu}{r}\right)^2};$$

$$k_\varphi = \frac{\nu}{r}.$$

2. GREEN'S FUNCTION

Let us determine the Green's function $\Gamma = \Gamma(r, \varphi, r', \varphi')$ of the problem, i.e., the function with the following properties.

1. It satisfies Equation (1), with the exception of the point (r', φ') where the source is located.
2. At the point (r', φ') it has a singularity corresponding to the two-dimensional δ function $-4\pi\delta(x - x', y - y')$, or in other words, near the source we have

$$\Gamma = i\pi H_0^{(1)}(k\rho) + \Gamma', \tag{12}$$

where ρ is the distance between the source and the point of observation; Γ' is a finite function. When $k\rho \ll 1$ we have $i\pi H_0^{(1)}(k\rho) \approx -2 \ln k\rho$.

3. As $|\varphi - \varphi'| \to \infty$ the function satisfies the radiation condition, i.e., it is a wave traveling from the source.
4. On the surface r = a of the cylinder it satisfies the boundary condition (3) or (2)

Further we use in lieu of condition (12) the equivalent condition

$$\left.\frac{\partial\Gamma}{\partial\varphi}\right|_{\varphi=\varphi'+0}-\left.\frac{\partial\Gamma}{\partial\varphi}\right|_{\varphi=\varphi'-0}=-4\pi r'\delta(r-r'). \tag{13}$$

We seek the function Γ in the form

$$\Gamma=\sum_s A_s J_{\nu_s}(kr)\,e^{i\nu_s\varphi}e^{-i\nu_s\varphi'},\quad \varphi>\varphi',$$

$$\Gamma=\sum_s A_s J_{\nu_s}(kr)\,e^{-i\nu_s\varphi}e^{i\nu_s\varphi'},\quad \varphi<\varphi', \tag{14}$$

where A_s are the sought coefficients; n_s are roots of (7) or (8).

The series (14) satisfy the properties 1, 2, and 4. The coefficients A_s will be determined from condition (13). Substituting (14) in (13)

$$\sum_s \frac{A_s}{r} J_{\nu_s}(kr)\,2i\nu_s=-4\pi\delta(r-r')\frac{r'}{r},$$

multiplying both halves by J_{ν_s}, integrating from 0 to ka, and using the well-known relation

$$\int\frac{J_nJ_m}{x}\,dx=\frac{x\,[J_mJ'_n-J_nJ'_m]}{n^2-m^2}$$

along with the boundary condition (7) or (8), we obtain

$$A_s=i2\pi\,\frac{J_{\nu_s}(kr')}{\nu_sN_s}, \tag{15}$$

where

$$N_s=-\frac{ka}{2\nu_s}\frac{d}{d\nu}J_\nu\,\frac{d}{d(kr)}J_\nu\Big|_{\substack{\nu=\nu_s,\\ r=a}} \tag{16}$$

for boundary condition (7) and

$$N_s=\frac{ka}{2\nu_s}J_\nu(ka)\,\frac{d^2}{d\nu\,d(kr)}J_0\Big|_{\substack{r=\nu_s\\ r=a}} \tag{17}$$

for boundary condition (8).

The series

$$\Gamma=-\frac{4\pi i}{ka}\sum_s\frac{J_{\nu_s}(kr)\,J_{\nu_s}(kr')}{\dfrac{dJ_{\nu_s}(ka)}{d\nu}\,\dfrac{dJ_{\nu_s}(ka)}{d(kr)}}\,e^{i\nu_s|\varphi-\varphi'|}, \tag{18a}$$

$$\Gamma=\frac{4\pi i}{ka}\sum_s\frac{J_{\nu_s}(kr)\,J_{\nu_s}(kr')}{J_{\nu_s}(ka)\,\dfrac{d^2}{d\nu\,d(kr)}J_{\nu_s}(ka)}\,e^{i\nu_s|\varphi-\varphi'|} \tag{18b}$$

satisfies all the conditions imposed and is the solution of the problem for boundary conditions (2) and (3).

Let us place, for the case of boundary condition (8), the source on the wall ($r'=a$, $\varphi'=0$) and let us regard (18b) as a sum of residues. The Green's function can be rewritten in the form of an integral along the contour C_0 (Figure 1), including all the roots of (7) or

(8), located in the first quadrant of the complex plane,

$$\Gamma = \frac{2}{ka} \int\limits_{C_0} \frac{J_\nu(kr)}{J'_\nu(ka)} e^{i\nu|\varphi|} d\nu. \tag{19}$$

3. ASYMPTOTIC EXPRESSION FOR Γ

The field can be calculated directly from formulas (18) only for relatively small ka. To calculate the field away from the source ($k\rho \gg 1$) it is necessary to take into account at least all the propagating waves. As will be shown below, when $ka \ll 1$ part of the normal waves, for which $kr > \nu$, can be replaced by their rays.

When $|\varphi - \varphi| < \pi$, the number of rays is much less than the number of normal waves. The introduction of rays clarifies the physical picture of the field and greatly facilitates the calculation.

However, the normal waves, for which $\nu_s \simeq ks$, cannot be represented in the form of any rays whatever and must be interpreted as whispering-gallery waves.

We shall consider further the region of points of observation away from the point of excitation ($k\rho \gg 1$), but located on the first sheet ($|\varphi - \varphi'| < \pi$).

To change over to the ray representation, we transform the contour C_0 into the contour C_1 (Figure 1), which covers only part of the real roots and then goes upward to infinity.

$$0 < \nu < ka - (ka)^{1/3} \tag{20}$$

The integral over the region remaining beyond C_1 will be represented in the form of a sum of residues. This procedure, by which the resultant field is represented in the form of a sum of rays and normal waves, is not unique. It remains valid, however, for any number of separated normal waves. Taking the foregoing into account, we have

$$\Gamma = \frac{2}{ka} \int\limits_{C_1} \frac{J_\nu(kr)}{J'_\nu(ka)} e^{i\nu|\varphi|} d\nu + \frac{4\pi i}{ka} \sum_s \frac{J_{\nu_s}(kr)}{\frac{d^2}{d\nu d(kr)} J_{\nu_s}(ka)} e^{i\nu_s|\varphi|}, \tag{21}$$

and the sum (21) is taken over all the roots of (8) which do not satisfy (20).

In the calculation of the integral (21) we employ the method of stationary phase. For the stationary-phase points $\tilde{\nu}$, satisfying the condition

$$0 < \tilde{\nu} < kr - (kr)^{1/3}, \tag{22}$$

the calculations yield the rays of geometrical optics. The stationary-phase points in the region

$$kr - (kr)^{1/3} < \tilde{\nu} < ka - (ka)^{1/3} \tag{23}$$

correspond to the field near and beyond the caustics of the rays.

A. The rays of geometric optics. We shall first assume that condition (22) is valid over the entire contour C_1, and we can consequently use the Debye asymptotic expansion (10) for J_ν. We shall not, however, use in the calculations all the roots of the stationary-phase equation, but only those roots that lie inside the interval (22).

Substituting (10) into the integral of (21) and recognizing that

$$J'_\nu(x) = -\sqrt{\frac{2}{\pi}} \frac{\sqrt[4]{x^2 - \nu^2}}{x} \sin \xi,$$

we obtain

$$\frac{\cos \xi_r}{\cos \xi_a} = -i\left(e^{i\xi_r} + e^{-i\xi_r}\right) \sum_0^\infty e^{i(2n+1)\xi_a},$$

$$\frac{2}{ka} \int\limits_{C_1} \frac{J_\nu(kr)}{J'_\nu(kr)} e^{i\nu|\varphi|} d\nu = 2i \sum_0^\infty \int\limits_{C_1} \frac{e^{i\psi_n^+} + e^{i\psi_n^-}}{\sqrt[4]{(ka)^2 - \nu^2}\sqrt[4]{(kr)^2 - \nu^2}} d\nu, \tag{24}$$

where

$$\psi_n^\pm = \pm \xi_r + (2n+1)\xi_a + \nu\varphi;$$

$$\xi_r = \sqrt{(kr)^2 - \nu^2} - \nu \arccos \frac{\nu}{kr} - \frac{\pi}{4};$$

$$\xi_a = \sqrt{(ka)^2 - \nu^2} - \nu \arccos \frac{\nu}{ka} - \frac{\pi}{4}. \tag{25}$$

Each of the integrals in the right half of (24) will be evaluated by the stationary-phase method.

The stationary-phase points $\widetilde{\nu}$ are determined as the roots of the equation

$$\frac{d\psi_n^{\mp}}{d\nu} = \pm \arc\cos\frac{\widetilde{\nu}}{kr} + \varphi - (2n+1)\arc\cos\frac{\widetilde{\nu}}{ka} = 0 \tag{26}$$

or putting

$$\varphi_0 = \arc\cos\frac{\widetilde{\nu}}{ka}, \quad \Delta\varphi = \arc\cos\frac{\widetilde{\nu}}{kr}, \quad \varphi = (2n+1)\varphi_0 \mp \Delta\varphi. \tag{26a}$$

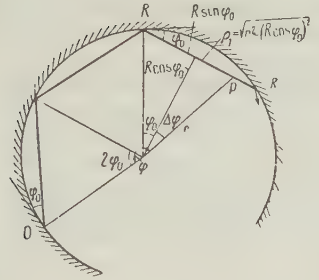


Figure 3. Construction of n-fold reflected ray; 0 — point of radiation, P — point of observation, R = a.

The geometric meaning of φ_0 and $\Delta\varphi$ is illustrated in Figure 3. A direct solution of (26) with respect to ν leads to an equation of degree 4n + 2 and is possible only when r = a.

We get around this difficulty first by assuming $\widetilde{\nu}$ specified and determining all the points of observation $r = r(\varphi)$ corresponding to the given stationary-phase point.

It follows from (26) that

$$r = \frac{a\cos\varphi_0}{\cos[\varphi - (2n+1)\varphi_0]}, \tag{27}$$

i.e., the geometric locus of $r(\varphi)$ is a straight line determined by the segment $a\cos\varphi_0$, turned by an angle $(2n+1)\varphi_0$.

It is easy to see that this line is a ray emerging towards the wall at an angle φ_0 and reflected n times. The case n = 0 corresponds to the primary wave of the source.

We now consider the envelope, the caustic of the family of rays experiencing n reflections.

From the parametric equations of the caustic

$$\frac{\partial r}{\partial\varphi_0} = 0 \quad \text{or} \quad \varphi = (2n+1)\varphi_0 - \arc\operatorname{tg}\frac{\operatorname{tg}\varphi_0}{2n+1} \tag{28}$$

and from (26) it follows that the point of tangency between the caustic and the ray (Figure 4) is located a distance $a\sin\varphi_0\left(1 - \frac{1}{2n+1}\right)$ from the point of reflection. For n = 0 the caustic is the source itself.

For $\varphi_0 < 1$ and $n > 1$ we can approximate the caustic equation in explicit form:*

$$r \simeq a\cos\frac{\varphi}{2n+1}. \tag{29}$$

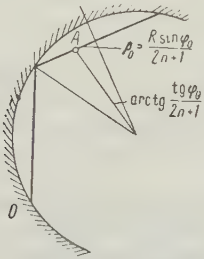

A
$\rho_0 = \frac{R\sin\varphi_0}{2n+1}$
$\arc\operatorname{tg}\frac{\operatorname{tg}\varphi_0}{2n+1}$
0

Figure 4. Construction of point of tangency and caustic: A — point of tangency; R = a.

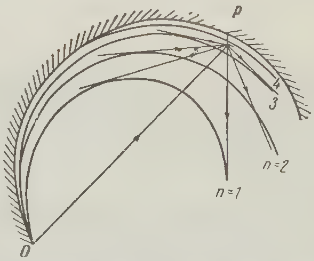


Figure 5. Caustics of n-fold reflected rays. Rays passing through the point of observation P should be tangent to the caustics. The total number of rays is determined by the number of caustics N, located between the point P and the center of the circle.

---

*It is interesting to note that in approximation (29) the caustics differ from each other only in a "scale" based on the angle φ.

Figure 5 shows the family of caustics for n = 1, 2, 3, 4.

We can now determine directly all the stationary points $\tilde{\nu}$ for each n for a specified point of observation (r,φ). They are determined in terms of the angles φ_0 for all 2N + 1 rays, which are tangent to the caustics and pass through the point (r,φ). For n > 0 there always exist two n-fold reflected rays, namely, one before tangency to the caustic (the term ψ^-) and the second after tangency to the caustic (the term ψ^+).

The total number of rays and, consequently, stationary-phase points in the sum (24) is limited to 2N + 1, i.e., it is determined by the number N of caustics lying between the point of observation and the center of the cylinder.

The closer the point of observation to the wall, the greater N. When r = 0 the solution of (26) is obtained in explicit form

$$\tilde{\nu} = ka \cos \frac{\varphi}{2m},$$

where

$$\begin{aligned} m &= n \quad \text{for} \quad \psi^+, \\ m &= n+1 \quad \text{for} \quad \psi^-. \end{aligned} \tag{30}$$

It follows from (30) that in the case when r = a, we have N = ∞, and the stationary-phase points condense near ka, which violates the initial approximation of J_ν for $n \gg 1$.

We now consider the amplitudes and phases of the individual rays. Writing

$$(\psi^{\mp})'' = \mp \frac{1}{\sqrt{(kr)^2 - \nu^2}} + (2n+1) \frac{1}{\sqrt{(ka)^2 + \nu^2}}, \tag{31}$$

substituting for $\tilde{\nu}$ the expression ka cos φ_0, and integrating (24), we obtain

$$\frac{2}{ka} \int \frac{J_\nu(kr)}{J'_\nu(ka)} e^{i\nu|\varphi|} d\nu = 2i\sqrt{\lambda} \sum_0^{2N+1} \frac{1}{\sqrt{(2n+1)|\rho_0 \mp \rho_1|}} \times$$
$$\times \exp\left\{i\left[k\left[(2n+1)a\sin\varphi_{0n} \mp \varrho_1\right] - \left(n + \frac{1}{2} \mp \frac{1}{2}\right)\frac{\pi}{2} + \frac{\pi}{4}\sin(\rho_1 \mp \rho_0)\right]\right\}, \tag{32}$$

where

$$\begin{aligned} \rho_0 &= \sqrt{r^2 - (a\cos\varphi_0)^2}; \\ \rho_1 &= \frac{a\sin\varphi_0}{2n+1}. \end{aligned} \tag{33}$$

The geometric meaning of ρ_0 and ρ_1 is clear from Figures 3 and 4. It follows from (32) that the rays satisfy the laws of geometric optics, i.e., a) the power flux is carried forward by the ray tube, b) the phase is proportional to the length of the path, c) on passing through the caustic the phase experiences a jump of $-\pi/2$. After n reflections the ray tube broadens by a factor 2n + 1.

B. Field near and beyond the caustic (point of observation does not lie on the wall (r < a)). It follows from the foregoing that some of the stationary-phase points lie in a region where the Debye expansion for $J_\nu(kr)$ is not valid, although it is still valid for $J'_\nu(ka)$. From the geometr cal point of view this means either that the point of observation is located near the caustic of some rays, or that some of the rays "fly by" between the point of observation and the wall. Whe as in the former case the laws of geometric optics yield values for the amplitudes that are too high, in the latter case they yield "zero" amplitudes. Actually, the diffraction "smears" both phenomena. Let us examine these phenomena in greater detail. If the point of observation moves along a ray (i.e., $\tilde{\nu}$ = const for one of the members of the sum), then at the center of the ray $\tilde{\nu}$ = kr and the Debye asymptotic approximation is inapplicable. In addition, at the point $\rho_0 = \rho_1$ where the ray is tangent to the caustic, we have $\psi''(\tilde{\nu}) = 0$, as follows from (31) and the integration carried out above is invalid.

Although the Debye asymptotic approximation is formally inapplicable near the center of the ray, it does yield the correct answer in the present approximation, for according to (32) the amplitude varies smoothly along the ray. The reason for this is that the product $\sqrt{(kr)^2 - \nu^2}\sqrt{\psi''}$ remains continuous, at the center of the beam.

Near the point of tangency between the beam and the caustic it is necessary to retain the cubic term in the expansion for ψ. This introduces an Airy function upon integration of (24). This approximation goes into the approximation of geometrical optics away from the caustic, and yields a finite value of the amplitude in the region of the caustic. The maximum values of the amplitudes are not on the caustic line, but are shifted somewhat closer to the cylinder boundary, i.e., on the two sides of the point of tangency between the beam and the caustic if the motion is along the beam. If the point of observation is displaced beyond the boundary of the caustic, closer to the center of the cylinder, it is necessary to use for the numerator of the integrand in (21) the Fok asymptotic approximation

$$J_\nu(kr) = \frac{1}{\sqrt{\pi}}\left(\frac{2}{kr}\right)^{1/3} v(t), \tag{34}$$

where v(t) is the Airy function;

$$\begin{gathered} t = \left(\frac{2}{kr}\right)^{1/3}(\nu - kr); \\ v(t) \sim |t|^{-1/4} \sin\left[\frac{2}{3}(-t)^{3/2} + \frac{\pi}{4}\right], \quad t < -1; \\ v(t) \sim \frac{1}{2} t^{-1/4} \exp\left[-\frac{2}{3} t^{3/2}\right], \quad t > 1. \end{gathered} \tag{35}$$

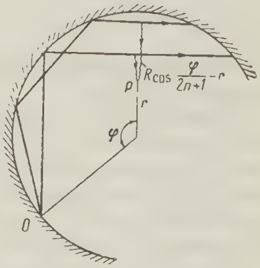


Figure 6. Contribution made to the field at the point of observation P by the rays passing near P and the center of the circle.

To calculate the influence of the rays located between the point of observation and the wall it is necessary to use approximation (34) for $J_\nu(kr)$. The position of the stationary-phase points

$$\varphi_{0m} = \frac{\varphi}{2m+1}$$

shows (see Figure 6) that the main contribution to the field is made by the nearest m-fold reflected ray, traveling between the point of observation and the wall of the cylinder. The amplitude of this field depends exponentially on the distance between the ray and the point of observation.

The propagation of the wave along this section (from the ray to the point of observation) takes place with infinite phase velocity (as in the case of the earth-surface waveguide when the point of observation is located above the waveguide channel [8]).

4. WHISPERING GALLERY WAVE

As shown above, the integral over the contour C_1 gives the field of rays of geometrical optics and the field in the vicinity of the caustic. It was assumed everywhere in the calculations that $\nu < ka - (ka)^{1/3}$, i.e., some of the roots of (8) were encircled by the contour C_1, and the integral was replaced by a sum of residues.

We shall now transform this sum in (21) and analyze the contributions of these waves.

When $\nu \sim ka$ and $\nu \sim kr$ we can employ for $J_\nu(kr)$ and Fok's asymptotic expression [7], according to which

$$J_\nu(kr) = \frac{1}{\sqrt{\pi}}\left(\frac{2}{kr}\right)^{1/3} v(t), \tag{36}$$

$$\frac{d^2 J_\nu(kr)}{d\nu\, d(kr)} = -\frac{1}{\sqrt{\pi}}\frac{2}{kr}\, t v(t). \tag{37}$$

Substituting (36) and (37) into the sum (21), we obtain

$$\frac{4\pi i}{ka}\sum_s \frac{J_{\nu_s}(kr)}{\frac{d^2}{d\nu d(kr)}J_\nu(ka)}\, e^{i\nu_s\varphi} = -2\pi i\left(\frac{2}{ka}\right)^{1/3} m \sum_s \frac{v\,[m\,(t_s+y)]}{v\,(t_s)\,t_s}\, e^{ip_s\varphi}, \tag{38}$$

where

$$m = \left(\frac{a}{r}\right)^{1/3}; \quad y = k\,(a-r)\left(\frac{2}{ka}\right)^{1/3}; \qquad p_s = ka - \left(\frac{ka}{2}\right)^{1/3} t_s;$$

t_s is the s-th root of the equation $v'(t) = 0$;

$$\begin{gathered} t_0 = -1.01; \;\; t_2 = -4.82; \\ t_1 = -3.25; \; t_s \simeq -\left\{\frac{3}{2}\left[(2s+1)\frac{\pi}{2} - \frac{\pi}{4}\right]\right\}^{2/3}. \end{gathered} \tag{39}$$

It follows from (38) and (39) that $t_s < 0$, and the quantity $t_s + y$ can be both less than and greater than zero.

When $t_s + y < 0$, the proper wave has an oscillating character, while when $t_s > 0$, i.e., away from the wall, it is damped. In analogy with the field of rays bounding between the point of observation and the wall, the first normal wave makes weak contributions to the field away from the wall. When $t_s + y < 0$, the field is produced both by the usual wave and by the normal waves. Since the functions v(t) have been tabulated in detail [8], calculations by means of formula (38) entail no great difficulty.

When r = a the sum (38) becomes much simpler, viz.,

$$2\pi i\left(\frac{2}{ka}\right)^{1/3}\sum_s \frac{e^{ip_s\varphi}}{t_s}. \tag{40}$$

It follows from the last expression that when $ka \gg 1$ there travels over the surface of the cylinder a wave whose amplitude decreases but little with increasing φ. The fundamental wave among this totality of normal waves is the lowest-order natural mode of the cylinder.

Thus, calculations confirm Rayleigh's point of view [3] that the anomalously weak damping of sound directly at the concave wall is due not to the focusing of the reflected waves, but to the presence of a specific wave hugging the concave wall. This wall is analogous to the wave in the earth's-surface waveguide [8].

CONCLUSION

The results of previously made calculations of diffraction on convex bodies shows that in the limit as $\lambda \to 0$, the field can be represented as a sum of "rays" satisfying the Fermat principle. Some of these rays are the ordinary rays of geometric optics, and some are diffraction rays [2], traveling over the surface of the body and attenuating exponentially in the shadow region.

The results of the present calculations show that the asymptotic representation of the field in the form of a sum of rays is valid also for diffraction on a concave surface. A characteristic of this case is the multiplicity of the rays of geometric optics. The equivalent of Keller's damped diffraction rays are the undamped whispering-gallery rays, traveling along the surface of the body. These rays are better regarded as being of the waveguide type, guided by the concave surface of the body.

The reduced rate of field attenuation is noticeable even when the point of observation moves away from the source at a constant distance from the wall greater than the thickness of the whispering-gallery wave layer. As φ is increased, the number of incoming rays is increased, as a result of the crossing of the caustic lines, and the field attenuates on the average more slowly than $\sqrt{a\varphi}$.

In conclusion, I consider it my pleasant duty to thank L.A. Vaynshteyn and A.A. Fedorov for interest and valuable advice as this paper was being written.

REFERENCES

1. D. Mentzer, Diffraction and Scattering of Radio Waves (Translated from the English), Soviet Radio Press, 1958.
2. L.A. Vaynshteyn and A.A. Fedorov, Scattering of Plane and Cylindrical Waves by an Elliptical Cylinder, and the Concept of Diffraction Rays. Radiotekhnika i elektronika, 1961, 6, 1, 31.
3. Rayleigh, Theory of Sound (Russian Translation), II, Section 287, GTI, 1940.
4. P.E. Krasnushkin, Metod normal'nykh voln v primenenii k problemye dal'nikh radiocbyazyey, (The Method of Normal Waves as Applied to the Problem of Long-Distance Radio Communication, Izd MGU, 1947.
5. G.A. Grinberg, N.N. Lebedev, N.P. Skal'skaya, Ya. S. Uflyand. The Wave Problem for a Parabolic Cylinder, Doklady AN SSSR, 1954, 95, 5, 951.
6. V.A. Fok, Theory of Diffraction by a Paraboloid of Revolution, Sb Difraktsiya elektromagnitnykh voln na nekotorykh telakh vrashcheniya (Collection, Diffraction of Electromagnetic Waves on Certain Bodies of Revolution), Izd. Sovyetskoye radio, 1957.
7. V.A. Fok, Difraktsiya radiovoln vokrug zemnoy poverkhnosti, (Diffraction of Radio Waves Around the Earth's Surface), Izd, AN SSSR, 1946.
8. L.M. Brekhovskikh, Waves in Layered Media, Izd. AN SSSR, 1959.

Submitted to the editors, 18 October 1960

SUPPRESSION OF RESONANCE PHENOMENA IN MULTIMODE WAVEGUIDES

R.B. Vaganov and V.V. Meriakri

Specifications are listed for filters for the suppression of undesirable modes, based on the premise that resonance phenomena are to be supressed. A method is described for the measurement of the parameters that characterize the resonant parts of the system; a method is proposed for eliminating the resonant effects, using strongly coupled directional couplers; the results of experiments on a specific filter model are reported.

INTRODUCTION

Waveguides of increased cross section, in which the operating modes can be accompanied by undesirable (parasitic) modes, are finding more frequent use in antenna techniques.

---

*From the paper by R.B. Vaganov, Yu.M. Isayenko,V.V. Meriakri, "Resonance Phenomena in Multimode Waveguides and How to Combat Them," delivered at the Session of the A.S. Popov Scientific-Technical Society for Radio Engineering and Electric Communication, 19 May 1960.

The greatest interest is attached to round waveguides in the H_{01} mode, in which attenuations on the order of several decibels per kilometer are attainable [1] . Undesired modes, brought about by inhomogeneities of the system, usually have a small amplitude; being resonant, however, they can cause large added losses * in the transmitted fundamental mode [2] , and also increase the standing wave ratio (SWR) of the resonating portions and produce overvoltages in them.

A classic example of an inhomogeneity that gives rise to parasitic waves is a smooth transition piece between two regular waveguides. When the fundamental wave travels through this transition, a parasitic wave is produced starting with a certain (critical) cross section and then propagates in the waveguide independently of the fundamental wave. A waveguide section capable of supporting two modes, situated between two such transitions, can under certain conditions resonate with the parasitic wave.

In reference [2] , the coefficient of transmission of such a section, of length L (Figure 1) is calculated by representing the transition in the form of a three-port with scattering matrix S. The transition is assumed lossless and matched on the input side. Under these conditions all the elements Γ_{mn} of the scattering matrix are expressed in terms of the coefficient of transformation of the fundamental wave into the parasitic wave, Γ_{02}. The expression derived in [2] for the transmission coefficient is

$$|D| = \frac{1}{1-\Gamma^2(1-e^{-j\varphi})}\left\{1-\frac{2\Gamma^2(1-\Gamma^2)(1-\cos\varphi)}{1-e^{-j2\psi_2}[1-\Gamma^2(1-e^{-j\varphi})]^2}\right\}. \tag{1}$$

Here $\Gamma = |\Gamma_{02}|$; $\varphi = \psi_1 - \psi_2$; $\psi_1 = (\beta_1 - j\alpha_1)\ L - \varphi_{11}$; $\psi_2 = (\beta_2 - j\alpha_2)\ L - \varphi_{22}$; β is the propagation constant; α is the attenuation factor in the two-mode waveguide; $\varphi_{11} = \arg \Gamma_{11}$ $\varphi_{22} = = \arg \Gamma_{22}$.

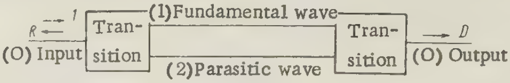


Figure 1. Diagram of resonant section.

Expressions for the remaining quantities in the cavity can be derived in similar fashion. We give here an expression for the coefficient of reflection from the cavity

$$|R| = \Gamma^2(1-\Gamma^2)\left|\frac{(1-e^{-j\varphi})e^{-j\psi_2}}{1-e^{-j2\psi_2}[1-\Gamma^2(1-e^{-j\varphi})]^2}\right|. \tag{2}$$

which is of interest to us.

A manifestation of the resonance is that when the conditions

$$\beta_2 L - \varphi_{22} = \pi p, \quad p = 1, 2, 3, \ldots \tag{3}$$

are satisfied, an absorption line is encountered in the signal passing through the portion of the channel. The losses introduced into the channel of the fundamental wave become maximal, and reflection from the resonating portion of the channel greatly increases, if the electric length of the resonator is such that the condition

$$(\beta_2 L - \varphi_{22}) - (\beta_1 L - \varphi_{11}) = (2q-1)\pi, \qquad q = 1, 2, 3, \ \ldots \tag{4}$$

is satisfied (or almost satisfied), in addition to condition (3) .

Let us examine a case of practical importance, that of small damping. If the resonating section is assumed to be sufficiently short and not to contain any elements that introduce any losses in the principal or parasitic waves, then $\alpha_1 L \ll 1$ and $\alpha_2 L \ll 1$. Under these conditions expressions (1) and (2) simplify to

---

*The term "fundamental wave" is used throughout this article to describe the operating mode.

$$|D| = \frac{1-\Gamma^2\left(1+\frac{\alpha_1}{\alpha_2}\right)}{1+\frac{2\Gamma^2(1-\Gamma^2)}{\alpha_2 L}} \simeq \frac{1}{1+\frac{2\Gamma^2}{\alpha_2 L}} \quad \text{for } \Gamma^2 \ll 1, \tag{5}$$

$$|R| = \frac{1}{1+\frac{\alpha_2 L}{2\Gamma^2(1-\Gamma^2)}} \simeq \frac{1}{1+\frac{\alpha_1 L}{2\Gamma^2}} \quad \text{for } \Gamma^2 \ll 1. \tag{6}$$

In the absence of losses in the parasitic mode ($\alpha_2 L = 0$), the fundamental wave does not pass at all through the resonating portion and is completely reflected. To reduce the direct losses and to decrease the SWR, which equals $(1 + |R|)/(1 - |R|)$, it is necessary to have $\alpha_2 L \gg 2\Gamma^2(1 - \Gamma^2)$. Another undesirable effect may be that the increase in the field intensity of the parasitic wave can produce a breakdown in the resonator. The maximum value of this intensity is $\frac{1}{\Gamma + \frac{\alpha_2 L}{2\Gamma}}$ in the case of maximum resonance (3) and (4) and of low losses; to reduce the overvoltage it is necessary to satisfy the condition $\alpha_2 L \gg 2\Gamma(1 - \Gamma)$.

Obviously, to eliminate these phenomena one can either develop transitions with minimum transformation coefficient, or, with the transition specified, introduce additional damping for the parasitic wave.

Yu. M. Isayenko has demonstrated [3] that the only way to prevent a resonator from responding to an application of short pulses to the inhomogeneities of the channel is to increase the value of $\alpha_2 L$.

1. METHOD OF MEASURING THE PARAMETERS THAT CHARACTERIZE THE RESONANT PORTIONS OF THE CHANNEL

In order to ascertain the extent to which a given channel is endangered by resonances at the parasitic wavelength, we must know the transformation coefficients Γ^2 of the transition pieces used in the channel, and the attenuation $\alpha_2 L$ of the parasitic wave. The calculation of these quantities for a real system that includes cones and other inhomogeneities is quite cumbersome, and it is thus desirable to have a simple method of measuring small values of Γ^2 and $\alpha_2 L$. In addition, it is necessary to be able to measure large values of $\alpha_2 L$, such as filters for the parasitic waves can introduce into the waveguide channel.

The first of these problems is simplest to solve by "transmission" measurements. In this case, we determine the transmission coefficients $|D|$ through a portion of the waveguide channel between two transitions at maximum resonance, and the change in resonator length δL necessary to make the power transmission coefficient equal to

$$|D_{1/2}|^2 = \frac{1+|D|^2}{2}. \tag{7}$$

We shall show that to determine Γ^2 and small attenuations $\alpha_2 L$ it is sufficient to know $|D|$ and δL. Indeed, we can obtain from (1) (making use of the smallness of $\alpha_2 L$) and from condition (7)

$$\delta(\beta_2 L - \varphi_{22}) = 2\Gamma^2 + \alpha_2 L. \tag{8}$$

We note that β_2 and φ_{22} depend only on the frequency, whereas the measurements involve only a change in the resonator length L, and, therefore, $\delta(\beta_2 L - \varphi_{22}) = \beta_2 \delta L$. From (5) and (8) we can readily obtain expressions for the conversion coefficient

$$\Gamma^2 = \frac{1}{2}\beta_2 \delta L(1 - |D|) \tag{9}$$

and for the damping of the parasitic wave

$$\alpha_2 L = \beta_2 \delta L |D|. \tag{10}$$

Knowing Γ^2 and $\alpha_2 L$ we can calculate both the coefficient of reflection from the resonator, and the maximum intensity of the parasitic wave in resonance.

It must be stipulated that this method can be used only if two identical transitions are available. Let us consider now what advantages this method has over the previously described "reflection' method [4, 5], where one measures the reflection from a waveguide system consisting of a transition and a multimode waveguide with reflecting piston.

a) The resonant transmission and reflection coefficients in circuits based on transmission and on reflection depend in different fashion on the electric length of the resonator, and the phase relationships between the fundamental and parasitic waves become essential here. In a circuit based on transmission, the resonance vanishes when the condition

$$(\beta_2 L - \varphi_{22}) - (\beta_1 L - \varphi_{11}) = 2q\pi, \quad q = 1, 2, 3, \ldots$$

is satisfied (missing resonance). The envelope of the resonances between two "missing" resonances is shown in Figure 2a. The form of the envelope is independent of the ratio

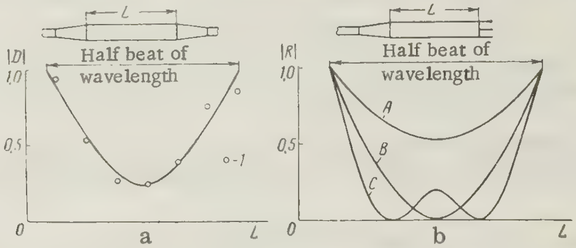


Figure 2. Envelopes of coefficients of transmission and coefficients of reflection in resonance: a — Measurement by transmission, 1 — experimental point; b — measurement by reflection

$A - \alpha_2 L/\Gamma^2 > 1$, $B - \alpha_2 L/\Gamma^2 = 1$, $C - \alpha_2 L/\Gamma^2 < 1$

$\alpha_2 l/\Gamma^2$, the magnitude of which changes only the maximum depth of the resonance. The "central" resonance, located between the missing resonances and representing the greatest interest, satisfies relation (4) and will be deepest in the case of measurement by transmission, so that its determination is made easiest. In the measurement by reflection, the situation is incomparably more complicated. The envelopes of the resonances are shown for this case in Figure 2b. Their form depends essentially on the value of $\alpha_2 L/\Gamma^2$ (see reference [4]). The central resonance can be determined only when all the resonances are known, at least within one-half of the beat wavelength.

b) It is seen from (9) and (10) that Γ^2 and $\alpha_2 L$ are uniquely defined in terms of $|D|$ and δL, whereas, in the reflection measurements it is necessary to establish first which is greater, Γ^2 or $\alpha_2 L$. Actually, in the latter case we have

$$\left.\begin{array}{c}\Gamma^2\\ \text{or}\\ \alpha_2 L\end{array}\right\} = \frac{1}{2}\beta_2\delta L\,(1 \pm R),$$

and the plus sign pertains to the greater of the two quantities Γ^2 or $\alpha_2 L$, while the minus sign pertains to the smaller one.

c) The width of the resonance curve is in our case greater than in the reflection method, owing to the appearance of an additional coupling element in the resonator, namely the second transition. In this connection, measurement of δL for small values of Γ^2 and $\alpha_2 L$ becomes simpler.

d) Finally, an important advantage of the measurement by transmission, in its practical realization, is the lack of the long-line effect that occurs in the measurements of reflection from the resonator as a result of reflections from the exciter or from the attenuators, and also as a result of the finite directivity coefficient of the coupler.

We note that a general shortcoming of these methods is the difficulty in measuring Γ^2 if $\Gamma^2 \gg \alpha_2 L$. It becomes necessary to introduce additional attenuation so as to bring these quantities closer to each other.

To solve the second problem, i.e., to measure the considerable attenuation introduced by a filter into a parasitic wave, it becomes necessary to use the most general formula (1) for the transmission coefficient. We shall assume that the resonance is maximal; in addition, the filter attenuates predominantly the parasitic wave, so that $\alpha_2 L \gg \alpha_1 L$, and the formula becomes somewhat simpler:

$$D = \frac{1}{1-\Gamma^2(1+e^{\alpha_2 L})}\left\{1 - \frac{2\Gamma^2(1-\Gamma^2)(1-\mathrm{ch}\,\alpha_2 L)}{1-e^{-2\alpha_2 L}[1-\Gamma^2(1+e^{\alpha_2 L})]^2}\right\}. \qquad (11)$$

We first measure the transformation coefficient Γ^2 of the transitions, using the procedure described above, and then the attenuation is introduced into the parasitic wave. By measuring the transmission coefficient $|D|$, we determine $\alpha_2 L$ from the family of curves $|D| = f(\alpha_2 L\,\Gamma^2)$, (Figure 3), which is calculated with the aid of formula (11).

Thus, even when the parasitic wave is appreciably attenuated, it is possible to determine the basic parameters of the resonator, so that both the SWR and the overvoltage in the waveguide can be determined from the general formulas.

2. STRONGLY COUPLED DIRECTIONAL COUPLER FOR THE SUPPRESSION OF RESONANCES AT THE H_{0n} MODE

To introduce additional damping of the parasitic waves, elements that give rise to thermal or radiation losses are usually placed in those regions of the waveguide where the field intensity (E or H) of the fundamental wave is close to zero, and the corresponding intensities of the filtered parasitic wave are finite. It is thus impossible to suppress in this fashion waves with a field configuration similar to that of the fundamental wave, without noticeably attenuating the latter. The literature contains a theoretical analysis of two types of H_{02} mode filters using a round waveguide so modified that the H_{02} mode is in the rejection band and the H_{01} mode in the transmission band. Symmetrical inhomogeneities, such as dielectric disks [6] , or else metallic diaphragms are introduced into the waveguide for this purpose. There is still no information, however, on any practical realization of filters for H_{0n} modes ($n \geqslant 2$) .

To suppress waves with a field structure similar to that of the fundamental wave, we use in the present investigation the principle of selection by phase velocities, i.e., we make use of the difference in phase velocity between the fundamental and parasitic waves. Filters based on this principle are made in the form of strongly coupled directional couplers, similar to Miller's well known diffraction converters [8] . Such a coupler, diverting the greater part of the energy of the parasitic wave from the main waveguide into an auxiliary waveguide, increases the value of $\alpha_2 L$ of the resonating wave, and consequently reduces the resonant effects.* We note than when a coupler is used as a filter for the suppression of resonance phenomena, the requirements imposed on the necessary coupling are less stringent than in the classical application of a directional coupler as a mode converter. In addition, the coupler need not have very high directivity. Coupling coefficients on the order of 0.5 — 1 db can at present be attained quite readily [9, 10] .

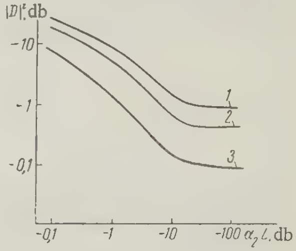


Figure 3. Dependence of the transmission coefficient on the ohmic losses of the parasitic wave.
1 - $\Gamma^2 = 1 \cdot 10^{-1}$; 2 — $\Gamma^2 = 5 \cdot 10^{-2}$; 3 — $\Gamma^2 = 1 \cdot 10^{-2}$.

Let us describe briefly the procedure used in the design of the coupler filter. To obtain maximum diversion of energy into the auxiliary waveguide, it is necessary to satisfy the following conditions [8] :

$$c_N' = \frac{\pi}{2}, \quad \beta_1' = \beta_2', \qquad (12)$$

---

*This method of suppression of resonant phenomena was proposed by Yu. M. Isaenko and V.V. Meriakri in 1958.

where c_N is the voltage coupling coefficient of the interacting wave through N apertures; β'_1 and β'_2 are the phase constants of the coupled waves.

In the determination of c_N, we shall confine ourselves to an analysis of coupling between two waves only, β_1 and β_2 (one each in the main and auxiliary waveguides), something that is always applicable if the selectivity of the coupler is sufficiently high. Let the coupled waves have different phase constants β'_1 and β'_2. Then the coefficient of excitation B_2 of the secondary wave can be represented, in accordance with [11], in the form

$$B_2 = -\frac{\int_S [\mathbf{E}\mathbf{h}_2]\, ds}{2k\beta' g^2} \tag{13}$$

where the integral is taken over the surface S of the coupling apertures; $k = 2\pi/\lambda$; $g = 2\pi/\lambda_{cr}$; h_2 is the orthonormalized auxiliary function of the magnetic field of the wave Π_2; E is the intensity of the electric field of wave Π_1 at the apertures λ is the wavelength in free space; λ_{cr} is the cutoof wavelength of the waveguide.

To obtain the amplitude B_2, it is necessary to know the distribution of E on the coupling element. Such a problem has been solved for a coupling between a round and rectangular waveguide through a round hole of radius $\rho \ll \lambda$, [12]. However, for example, for H_{0n} modes the most effective couplers are those having longitudinal slots, which yield a maximum coupling coefficient per unit length. A direct solution of such a problem involves great difficulties. In this case we can estimate the coupling coefficient by using the procedure of reference [12], taking into account the relation between a slot of length l and an equivalent round hole of radius ρ:

$$\frac{4}{3}\rho^3 = Ml^3, \tag{14}$$

where M is the magnetic polarizability of the slot [13].

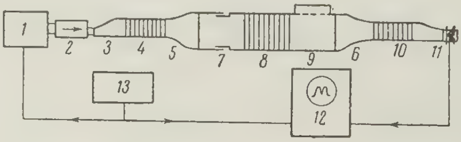


Figure 4. Block diagram of measuring apparatus.

Having determined the amplitude B_2 of the excited wave (13), we can readily determine an expression for the coupling c_N for coupling through N identical slots, after first calculating the power of the Π_2 wave. For the case of coupling between $H_{0n}^{\bigcirc}$ and $H_{10}^{\square}$ waves* this calculation leads to the equation

$$c_N = \sqrt{\frac{\pi}{2}}\,\frac{\mu_{0n} N M l^3}{\beta_{0n}\sqrt{a^3 b r^2}}\, e^{j\beta_{\text{щ}}\tau}. \tag{15}$$

Here μ_{0n} is the n-th root of the equation $J'_0(x) = 0$; r is the radius of the round waveguide; a and b are the larger and smaller dimensions of the rectangular waveguide, respectively; τ is the thickness of the slot wall; β_{sl} is the phase constant of the Π_1 wave in the slot. Strictly speaking expressions (14) and (15) are valid only when ρ and l are much smaller than λ. At large values of ρ and l, the accuracy of the calculation decreases. However, in reference [9] the expression (14) was used to calculate the coupling between the $H_{11}^{\bigcirc}$ and $H_{10}^{\square}$ modes through long slots ($l > \lambda$), and the agreement between the experimental and theoretical results was satisfactory. Coupling between resonant slots ($l = n\lambda/2$, $n = 1, 2, 3, \ldots$), investigated in reference [10], is important only for narrow band directional couplers for individual frequency channels, and is not considered here.

---

*The symbols ○ and ▢ denote a round and rectangular waveguide, respectively.

The maximum transfer of energy from the wave II_1 to the wave II_2 as indicated in reference [12], takes place when $c_N = \pi/2$, and complete transformation is attained in the case when the propagation constants of these waves are equal. The presence of coupling changes the phase constants, compared with the phase constants β_1 and β_2 in uncoupled waveguides of like dimensions. The "perturbed" phase constants β'_1 and β'_2 have the following form [8]:

$$\beta_1' = \beta_1 + \frac{c_{11}}{d},$$
$$\beta_2' = \beta_2 + \frac{c_{22}}{d}, \qquad (16)$$

where c_{11} and c_{22} are the coefficients of coupling through one hole between two identical waves in identical coupled waveguides; d is the distance between neighboring holes. For the $H^{\bigcirc}_{0n} \to H^{\square}_{10}$ coupler, c_{11} corresponds to a coupling between the $H^{\bigcirc}_{0n}$ and $H^{\bigcirc}_{0n}$ modes, and c_{22} corresponds to a coupling between $H^{\square}_{10}$ and $H^{\square}_{10}$.

It is easy to obtain, in analogy with (15),

$$c_{11} = \frac{1}{2\pi}\frac{\mu_{0n}^2 M l^3}{r^4 \beta_{0n}},$$
$$c_{22} = \frac{\pi M l^3}{a^3 b \beta_{10}}. \qquad (17)$$

Conditions (12) and the derived relations (15), (16) and (17) enable us to select the dimensions and the number of coupling elements, and also the dimension of the auxiliary waveguide.

3. EXPERIMENTAL RESULTS

Using the procedure described in Section 2, we designed and constructed a model of an H^0_{02} mode filter in a round waveguide of 25 mm diameter, intended for wavelengths on the order of 8 mm. The coupling system consisted of a series of slots (up to 25) measuring 7 by 1.5 mm. The coupling coefficient through one slot [see expression (15)] between the H^0_{02} and H^0_{10} modes was approximately - 19 db. A total of 16 holes was sufficient to satisfy conditions (12); a larger number of coupling elements was taken for the purpose of investigating the dependence of c_N on N. The filter was prepared galvanically, thus improving greatly the quality of the contact at the junction between the round and rectangular waveguides.

Measurements of the suppression of resonance phenomena at the $H^{\bigcirc}_{02}$ mode with the aid of the filter were carried out in accordance with the scheme illustrated in Figure 4.

Power was fed from microwave oscillator 1 through a decoupling attenuator 2 to the $H^{\square}_{10} \to H^{\bigcirc}_{01}$ transformer 3 with ring filter 4. The $H^{\bigcirc}_{01}$ wave travels from the left into transition 5 from the 12 diameter to the 25 mm diameter. In the transition, an $H^{\bigcirc}_{02}$ mode is excited and is capable of resonating between transition 5 and the analogous transition 6. A movable joint 7 is connected between the two transitions, so that the length of the resonator can be changed; also connected between the transmissions is a ring filter 8, intended for suppression of the asymmetrical waves, and the investigated $H^{\bigcirc}_{02}$ -mode filter 9. At the output of the resonator is a $H^{\square}_{10} \to H^{\bigcirc}_{01}$ converter with filter 10. The signal is fed from the detector pickup 11 to the input of oscillograph 12. The frequency of generator 1 is sinusoidally modulated by an audio generator 13, so that the oscillation zone of the klystron can be observed on the oscilloscope screen and the resonant dip in the zone is seen in the case of resonance at the parasitic wavelength. By changing the length of the resonator with the aid of moving joint 7, we obtain a maximum dip. Comparing the signal passing through the system at resonance with the magnitude of the zone in the absence of resonance, we determine the transmission coefficient at resonan

Transitions 5 and 6 had $\Gamma^2 = 0.05$. The losses introduced by the filter 9 in the $H^{\bigcirc}_{02}$ mode reached 10 db in the wavelength interval 8.0 — 8.3 mm, (Figure 5). The dependence of $\alpha_2 L$ on the number of open slots, some of which were covered with foil (Figure 6), was determined at an arbitrary frequency point. To equalize the phase velocities of the coupled waves, a tuning metal plate was inserted into the rectangular waveguide. In all cases, the losses introduced by the filter at the $H^{\bigcirc}_{01}$ mode were less than 0.2 db.

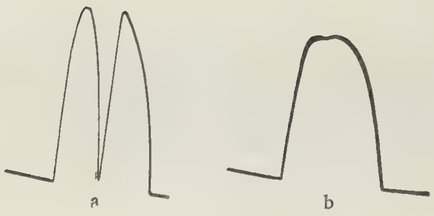


Figure 5. Losses at H_{01} wave in resonance to the H_{02} wave; a — Case of low losses at H_{02} wave; b — case of large losses at H_{02} wave.

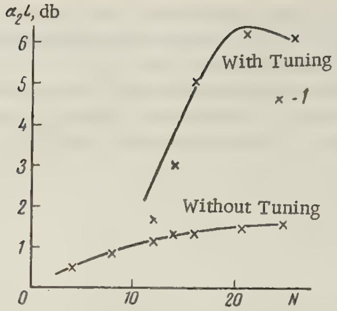


Figure 6. Dependence of the losses introduced at the H_{02} wave on the number of slots (1 — experimental points).

CONCLUSION

1. An analysis of the undesirable effects due to resonance at parasitic wavelengths in multimode waveguides makes it possible to calculate the increases in the losses and in the SWR of the fundamental wave, and also the overvoltages that result in feeders carrying several propagating waves. On the basis of the relations obtained, it is possible to determine in each specific case the optimum method of eliminating resonance effects.

2. Resonance phenomena can be experimentally investigated by the procedure proposed in this paper. A factor of great practical importance is the possibility of measuring the coefficients of wave excitation and the attenuation coefficients.

3. It was impossible hitherto to filter out waves whose field structure is similar to that of the working wave (in particular the H_{02} mode). The filtering method proposed, based on the use of strongly coupled directional couplers as filters, solve this problem to some extent. The simple procedure for approximate design of strongly coupled directional couplers, described in the paper, has been applied to the design of a filter for the H_{02}° mode in a waveguide with kr $\simeq$ 10. A directional coupler model, constructed with the use of galvanic technology, made it possible to produce filters whose insertion losses reach 10 db at the parasitic wave and do not exceed 0.2 db at the working wavelength.

4. On going to waveguides with larger diameters, where it becomes necessary to reduce the damping of the transmitted signal, the filter coupling coefficient is greatly reduced, in accordance with (18). To suppress resonance phenomena in lines having large kr, the filters therefore, must be placed near the critical section of the attenuated wave, i.e., in the region of minimum values of kr. If it is impossible to change over to a smaller diameter, it is expedient to use filters having a series of azimuthal channels, similar to the couplers proposed in reference [12]. At a specified parasitic-wave attenuation, this makes it possible to reduce greatly the length of the filter compared with the single-channel value. There is no need to make the channels symmetrical in phase, since the diverted power can be absorbed separately in each auxiliary waveguide.

The authors are very grateful to M. V. Persikov for valuable discussions and for interest in this investigation.

REFERENCES

1. Yu.I. Kaznacheyev, Radiotekhnika i elektronika, 1958, 3, 6, 731. Vestn. AN SSSR 1960, 2, 67.
2. A. King, E. Marcartili, Waveguide Transmission Line with Low Losses, Collection of Articles, translated from the English, I/, 1960, 374.
3. R.B. Vaganov, Yu.M. Isayenko, V.V. Meriakri, Paper delivered at the Session of the Scientific Technical A.S. Popov Society for Radio Engineering and Electric Communication, 19 May 1960.
4. Y. Klinger, Proceedings, Institute of Electrical Engineers, 1959, B106 Suppl. N13, 89.

5. R.B. Vaganov, Radiotekhnika i elektronika, 1960, 5, 5, 727.
6. M. Jouget, French patent, class $H_{01}b$, No. 1160875, 12 August 1958.
7. K.M. Arkhipova, V.A. Sudakov, Izv. AN SSSR, OTN (Energetika i avtomatika) (Izvest. of the Academy of Sciences, Division of Technical Sciences, Power and Automation,) 1960, 1, 1944.
8. S. Miller, Waveguide Transmission Line with Low Losses, Collection of articles, trans. from the English, IL 1960, 139.
9. R.W. Dawson, Bell System Technical Journal, 1957, 36, 2, 391.
10. V.V. Meriakri, Radiotekhnika i elektronika, 1959, 4, 11, 1929.
11. M.B. Zakson, Dokl AN SSSR, 1949, 66, 4, 637.
12. M.V. Persikov, Radiotekhnika i elektronika, 1957, 2, 1, 65.
13. S. Cohn, Proceedings IRE, 1951, 39, 11, 1417.

Institute of Radio Engineering and Electronics,
Academy of Sciences, USSR

Submitted to the editors, 25 August 1960

ELECTROMAGNETIC FIELD AND DISTRIBUTION OF CURRENT ALONG AN INFINITELY LONG INSULATED WIRE IN A CONDUCTING MEDIUM

P.P. Pavlov

We investigate the electromagnetic field and the current distribution along an infinitely long insulated wire in a conducting medium, with the excitation applied at a single point. It is demonstrated that the telegraphers equations can be used for an insulated wire situated in sea water and in moist earth, in the frequency range from $6 \cdot 10^3$ to $3 \cdot 10^7$ cps.

INTRODUCTION

When the basic parameters, such as the input resistance, the effective length, and directivity pattern are calculated, an antenna made in the form of an insulated wire of finite length is usually regarded as a segment of a lossy long line with open or short-circuited end or else loaded by a complex impedance of finite magnitude. Calculations for such lines are based on the use of the telegrapher's equations, but this approach to the problem calls for additional verification.

1. ELECTROMAGNETIC FIELD OF A CYLINDRICAL INSULATED WIRE SITUATED IN A CONDUCTING MEDIUM

In the analysis of the electromagnetic field produced by the current in an infinitely long wire, it is necessary to take into consideration three media (Figure 1): 1) the wire, with electromagnetic properties characterized by ϵ_1, μ_1, and σ_1, 2) a cylindrical cavity filled with insulating material, with parameters ϵ_2, μ_2, and σ_2 and 3) a conducting medium, with parameters ϵ_3. μ_3, and σ_3.

Assuming the time dependence of the field components and of the current to be in the form $e^{-i\omega t}$ the solution of the wave equation assumes the form

$$\Pi = \int_{-\infty}^{+\infty} A(\gamma) Z_0 (r \sqrt{k^2 - \gamma^2}) e^{i\gamma z} d\gamma. \tag{1}$$

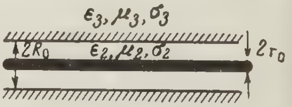


Figure 1

To simplify the problem, we shall assume that the specific electric conductivity of the wire is infinite. Under this assumption, the field inside the wire can be assumed equal to zero. As a result the problem of determining the field in three media reduces to a problem of finding the field in two media.

In view of the fact that waves are reflected from the interface between the dielectric and the conducting medium in a cylindrical cavity we write the solution of the wave equation inside the cylindrical cavity in the form

$$\Pi = \int_{-\infty}^{+\infty} [A_1 (\gamma) J_0 (v_2 r) + A_2 (\gamma) N_0 (v_2 r)] e^{i\gamma z} d\gamma, \tag{2}$$

where

$$v_2 = \sqrt{k_2^2 - \gamma^2}; \quad k_2 = \frac{2\pi}{\lambda} \sqrt{\varepsilon_{k2}'};$$

$\epsilon'k2$ is the relative dielectric constant of the insulating material. For an absorbing medium we have

$$\Pi = \int_{-\infty}^{+\infty} A_3 (\gamma) H_0^{(1)} (v_3 r) e^{i\gamma z} d\gamma, \tag{3}$$

where

$$v_3 = \sqrt{k_3^2 - \gamma^2}; \qquad k_3 = \frac{2\pi}{\lambda}\sqrt{\varepsilon'_{k3}};$$

ϵ'_{k3} is the relative dielectric constant of the conducting medium.

The components of the electromagnetic field are given by the following expressions:

a) in the cylindrical cavity

$$E_{r2} = -i\int_{-\infty}^{+\infty} \gamma v_2 \left[A_1(\gamma) J_1(v_2 r) + A_2(\gamma) N_1(v_2 r)\right] e^{i\gamma z} d\gamma, \tag{4}$$

$$E_{z2} = \int_{-\infty}^{+\infty} v_2^2 \left[A_1(\gamma) J_0(v_2 r) + A_2(\gamma) N_0(v_2 r)\right] e^{i\gamma z} d\gamma, \tag{5}$$

$$H_{\varphi 2} = -i\omega\varepsilon_{k2}\int_{-\infty}^{+\infty} v_2 \left[A_1(\gamma) J_1(v_2 r) + A_2(\gamma) N_1(v_2 r)\right] e^{i\gamma z} d\gamma, \tag{6}$$

b) in the absorbing medium

$$E_{r3} = -i\int_{-\infty}^{+\infty} \gamma v_3 A_3(\gamma) H_1^{(1)}(v_3 r) e^{i\gamma z} d\gamma, \tag{7}$$

$$E_{z3} = \int_{-\infty}^{+\infty} v_3^2 A_3(\gamma) H_0^{(1)}(v_3 r) e^{i\gamma z} d\gamma, \tag{8}$$

$$H_{\varphi 3} = -i\omega\varepsilon_{k3}\int_{-\infty}^{+\infty} v_3 A_3(\gamma) H_1^{(1)}(v_3 r) e^{i\gamma z} d\gamma. \tag{9}$$

The unknown functions $A_1(\gamma)$, $A_2(\gamma)$ and $A_3(\gamma)$ are determined from the boundary conditions which in our problem have the form

$$E_{z2} + E_{\text{ext}} = 0 \quad \text{for} \quad r = r_0, \tag{10}$$

$$\left.\begin{aligned} E_{z2} &= E_{z3} \\ H_{\varphi 2} &= H_{\varphi 3} \end{aligned}\right\} \quad \text{for} \quad r = R_0, \tag{11}$$

where E_{ext} is the intensity of the field of the extraneous forces acting on the wire surface.

As in references [1 — 3], we assume the intensity of the extraneous-force field to be non-vanishing only is the section of the wire $0 < z < a$. The emf of the external source is then

$$\mathscr{E}_0 = \int_0^a E_{\text{ext}} dz. \tag{12}$$

The intensity of the field E_{ext} due to the emf $\mathscr{E}_0$ of the external source can be represented in the form of an integral

$$E_{\text{ext}} = \frac{\mathscr{E}}{2\pi}\int_{-\infty}^{+\infty} \frac{\sin\gamma a}{\gamma a} e^{i\gamma z} d\gamma. \tag{13}$$

Introducting the notation

$$K(\gamma) = \frac{\mathscr{E}}{2\pi}\frac{\sin\gamma a}{\gamma a}, \tag{14}$$

we get

$$E_{\text{ext}} = \int_{-\infty}^{+\infty} K(\gamma)\, e^{i\gamma z}\, d\gamma. \tag{15}$$

Using the boundary conditions (10) amd (11), we obtain the following system of equations

$$J_0(v_2 r_0)\, A_1(\gamma) + N_0(v_2 r_0)\, A_2(\gamma) = -\frac{K(\gamma)}{v_2^2}, \tag{16}$$

$$J_0(v_2 R_0)\, A_1(\gamma) + N_0(v_2 R_0)\, A_2(\gamma) - \frac{v_3^2}{v_2^2} H_0^{(1)}(v_3 R_0)\, A_3(\gamma) = 0, \tag{17}$$

$$J_1(v_2 R_0)\, A_1(\gamma) + N_1(v_2 R_0)\, A_2(\gamma) - \frac{\varepsilon_{k3}}{\varepsilon_{k2}} \frac{v_3}{v_2} H_1^{(1)}(v_3 R_0)\, A_3(\gamma) = 0. \tag{18}$$

The coefficients $A_1(\gamma)$, $A_2(\gamma)$, and $A_3(\gamma)$ are respectively equal to

$$A_1(\gamma) = \frac{D_1}{D}, \quad A_2(\gamma) = \frac{D_2}{D}, \quad A_3(\gamma) = \frac{D_3}{D}.$$

The determinant of the system of equations is

$$D = \frac{v_3^2}{v_2^2} H_0^{(1)}(v_3 R_0)\, [J_0(v_2 r_0)\, N_1(v_2 R_0) - J_1(v_2 R_0)\, N_0(v_2 r_0)] -$$

$$- \frac{\varepsilon_{k3}}{\varepsilon_{k2}} \frac{v_3}{v_2} H_1^{(1)}(v_3 R_0)\, [J_0(v_2 r_0)\, N_0(v_2 R_0) - J_0(v_2 R_0)\, N_0(v_2 r_0)]. \tag{19}$$

The determinants D_1, D_2, and D_3, are, respectively,

$$D_1 = -\frac{K(\gamma)}{v_2^2} \left[\frac{v_3^2}{v_2^2} H_0^{(1)}(v_3 R_0)\, N_1(v_2 R_0) - \frac{\varepsilon_{k3}}{\varepsilon_{k2}} \frac{v_3}{v_2} H_1^{(1)}(v_3 R_0)\, N_0(v_2 R_0) \right], \tag{20}$$

$$D_2 = -\frac{K(\gamma)}{v_2^2} \left[\frac{\varepsilon_{k3}}{\varepsilon_{k2}} \frac{v_3}{v_2} H_1^{(1)}(v_3 R_0)\, J_0(v_2 R_0) - \frac{v_3^2}{v_2^2} H_0^{(1)}(v_3 R_0)\, J_1(v_2 R_0) \right], \tag{21}$$

$$D_3 = -\frac{K(\gamma)}{v_2^2} [J_0(v_2 R_0)\, N_1(v_2 R_0) - J_1(v_2 R_0)\, N_0(v_2 R_0)]. \tag{22}$$

2. DERIVATION OF EQUATION FOR CURRENT IN THE WIRE

In a cylindrical cavity filled with dielectric, the current I_z flowing in the wire is connected with the component $H_{\varphi 2}$ of the field by the relation

$$I_z = 2\pi r H_{\varphi 2}. \tag{23}$$

This relation is valid when $r \ll \lambda$. In our problem this inequality is satisfied, since the thickness of the cylindrical cavity, $R_0 - r_0$, is much smaller than the wavelength.

$$H_{\varphi 2} = -i\omega\varepsilon_{k2} \int_{-\infty}^{+\infty} \frac{D_{\varphi 2} K(\gamma)}{v_2 D} e^{i\gamma z} d\gamma. \tag{24}$$

Substituting the value of $K(\gamma)$ from (14) into (24), we obtain

$$H_{\varphi 2} = -\frac{i\omega\varepsilon_{k2}\mathscr{E}}{2\pi} \int_{-\infty}^{+\infty} \frac{D_{\varphi 2}}{v_2 D} \frac{\sin\gamma a}{\gamma a} e^{i\gamma z} d\gamma. \tag{25}$$

when the insulation layer is thin compared with the wavelength, we obtain by assuming $r = R_0$

$$D_{\varphi 2} = \frac{2\varepsilon_{k3} v_3}{\pi R_0 \varepsilon_{k2} v_2^2} H_1^{(1)}(v_3 R_0). \tag{26}$$

Table 1

frequency cps	$k_2\left(\varepsilon_2'=3,\ \sigma_2=0\right)$	Sea Water $\left(\varepsilon_3'=80,\ \sigma_3=4\,\frac{1}{\text{ohm}\cdot\text{m}}\right)$		Moist Earth $\left(\varepsilon_3'=10,\ \sigma_3=5\cdot10^{-2}\,\frac{1}{\text{ohm}\cdot\text{m}}\right)$		Dry Earth $\left(\varepsilon_3'=4,\ \sigma_3=10^{-3}\,\frac{1}{\text{ohm}\cdot\text{m}}\right)$	
		k_3	γ_{p_1}	k_3	γ_{p_1}	k_3	γ_{p_1}
$6\cdot10^3$	0.0002176559	0.307813+ + i0.307811	0.0005565+ + i0.0000304	0.03441557+ + i0.0344132	0.0006357+ + i0.00002668	0.004870181+ + i0.00486369	0.0006585152+ + i0.0000273291
$3\cdot10^4$	0.001088279	0.6882997+ + i0.688277	0.0026224+ + i0.00016143	0.0769658+ + i0.007694	0.0030388+ + i0.000148279	0.01091913+ + i0.01084658	0.003357081+ + i0.0001936987
$3\cdot10^5$	0.01088279	2.176922+ + i2.176197	0.023748+ + i0.0017835	0.2437526+ + i0.2429414	0.0286236+ + i0.0015743	0.03558002+ + i0.033287	0.02953152+ + i0.00610777
$3\cdot10^6$	0.01088279	6.894364+ + i6.871423	0.20989+ + i0.02024	0.7824601+ + i0.7568128	0.258337+ + i0.0219342	0.148761+ + i0.0796144	0.1506621+ + i0.05970374
$3\cdot10^7$	1.088279	22.13131+ + i21.4059	1.782413+ + i0.243579	2.866357+ + i2.065953	2.109522+ + i0.4364983	1.260147+ + i0.09398566	1.229379+ + i0.07996127

Substituting (26) in (25) and letting $a \to 0$, we obtain

$$H_{\varphi 2} = -i\frac{\omega\varepsilon_{k3}\mathscr{E}}{\pi^2 R_0}\int_{-\infty}^{+\infty}\frac{v_3}{v_2^3 D}H_1^{(1)}(v_3R_0)\,e^{i\gamma z}d\gamma. \qquad (27)$$

On the basis of (23) we have

$$I_z = -i\frac{2\omega\varepsilon_{k3}\mathscr{E}}{\pi}\int_{-\infty}^{+\infty}\frac{v_3}{v_2^3 D}H_1^{(1)}(v_3R_0)\,e^{i\gamma z}\,d\gamma. \qquad (28)$$

The functions $v_2 = \sqrt{k_2^2-\gamma^2}$ and $v_3 = \sqrt{k_3^2-\gamma^2}$ are double-valued functions in the plane of the complex variable γ. The branch points of these functions are $\gamma = \pm k_2$, $\gamma = \pm k_3$.

The integrand, which is a complicated function of the elementary functions v_2 and v_3 also has two branch points. In addition to these singularities, the integrand has a pole at $D = 0$, $\gamma = \gamma_{p1}$.

Setting (19) equal to zero, we obtain

$$\frac{J_0(v_2r_0)\,N_1(v_2R_0) - J_1(v_2R_0)\,N_0(v_2r_0)}{J_0(v_2r_0)\,N_0(v_2R_0) - J_0(v_2R_0)\,N_0(v_2r_0)} = \frac{\varepsilon_{k3}}{\varepsilon_{k2}}\,\frac{v_2}{v_3}\,\frac{H_1^{(1)}(v_3R_0)}{H_0^{(1)}(v_3R\,)}. \qquad (29)$$

In view of its complexity, Equation (29) cannot be solved with respect to the constant γ without some simplification.

Recognizing that when $r_0 < R_0 \ll 1$ we have $|v_2r_0| < |v_2R_0| \ll 1$ and $|v_3r_0| < |v_3R_0| \ll 1$, we replace the cylindrical functions by first terms of their series expansions. Discarding terms of higher order of smallness, we obtain

$$(k_0^2\varepsilon_{k3}' - \gamma^2)\left(\ln\frac{1{,}12}{R_0\sqrt{k_0^2\varepsilon_{k3}-\gamma^2}} + i\frac{\pi}{2}\right) = -\frac{\varepsilon_{k3}'}{\varepsilon_{k2}'}(k_0^2\varepsilon_{k2}' - \gamma^2)\ln\frac{R_0}{r_0}. \qquad (30)$$

Table 1 lists the values of the wave number k_3 and the separation constant γ_{p1} calculated for sea water, for moist earth, and for dry earth. The separation constant γ_{p1} has been calculated for $R_0/r_0 = 3$ ($R_0 = 6$ mm). Table 1 lists also the values of the wave number k_2, calculated for $\epsilon'_2 = 3$, $\sigma_2 = 0$. If we neglect the losses in the insulation of the wire ($\sigma_2 = 0$) then the branch points $\gamma = \pm k_3$ will lie on the real axis. The branch points $\gamma = \pm k_3$ lie near the pole for all media. As can be seen from Table 1, the distance between these points and the pole is much more pronounced than in the case of dry or moist earth.

After determining the singular points of the integrand, we can employ the standard procedure for obtaining an asymptotic expression for the integral (28).

If the integrand contains the factor $e^{i\gamma z}$ and the coordinate z is positive, we choose the branches of the functions v_2 and v_3 corresponding to the roots $\gamma = k_2$ and $\gamma = k_3$, and we also choose the root $\gamma = \gamma_{p1}$ for the equation $D = 0$. We shall henceforth omit the subscript γ.

As a result, we arrive at the integration contour shown in Figure 2. The radius of the arcs CR is assumed to be quite large ($R \to \infty$). We denote by L'_R the curve consisting of the double-edged cut, parallel to the positive β-axis, and the small circle around $k_3 = \gamma$. We analogously denote by L''_R the curve consisting of a double-edge cut, parallel to the positive β-axis, and the small circle around the point $k_2 = \gamma$.

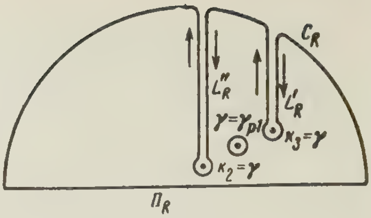


Figure 2.

On the basis of the residue theorem, the integral of (28) has a value

$$\mathcal{J} = \int\limits_{\Pi_R+C_R+L'_R+L''_R} \frac{v_3}{v_2^3 D} H_1^{(1)}(v_3R_0)\, e^{i\gamma z}\, d\gamma = 2\pi i R(z), \tag{31}$$

where R(z) is the residue of the integrand at the point $\gamma = \gamma_{p1}$.

When $R \to \infty$ and $z \neq 0$, we have by Jordan's lemma that the integral along the arc C_R tends to zero.

Calculation of the integral along the cut L'_R. We denote by $\mathcal{J}_{L'_R}$ the integral along the cut L''_R:

$$\mathcal{J}_{L'_R} = \int\limits_{L'_R} \frac{H_1^{(1)}(v_3R_0)\, e^{i\gamma z} d\gamma}{v_3 H_0^{(1)}(v_3R_0)\,\Phi_1(\gamma) - H_1^{(1)}(v_3R_0)\,\Phi_2(\gamma)}, \tag{32}$$

where

$$\Phi_1(\gamma) \equiv v_2\,[J_0(v_2r_0)\,N_1(v_2R_0) - J_1(v_2R_0)\,N_0(v_2r_0)]; \tag{33}$$

$$\Phi_2(\gamma) \equiv \frac{\varepsilon_{k3}}{\varepsilon_{k2}} v_2^2\,[J_0(v_2r_0)\,N_0(v_2R_0) - J_0(v_2R_0)\,N_0(v_2r_0)]. \tag{34}$$

Making the substitution

$$k_3 - \gamma = s, \quad d\gamma = -ds,$$

we obtain (Figure 3)

$$\int\limits_{L_R} = \underset{(1)}{\int\limits_{-iR}^{i\rho}} + \underset{(2)}{\int\limits_{-i\rho}^{-iR}} + \oint\limits_{C_\rho}.$$

As $\rho \to 0$, the integral along the circle C_ρ tends to zero. Along the first edge (see Figure 3) we have

$$s = \tau e^{i\frac{3\pi}{2}}, \quad v_{31} = \sqrt{\tau e^{i\frac{3\pi}{2}}(2k_3 + i\tau)}, \; d\gamma = -e^{i\frac{3\pi}{2}}\, d\tau.$$

Along the second edge

$$s = \tau e^{-i\frac{\pi}{2}}, \quad v_{32} = \sqrt{\tau e^{-i\frac{\pi}{2}}(2k_3 + i\tau)}, \quad d\gamma = -e^{-i\frac{\pi}{2}}\, d\tau.$$

We thus have

$$\mathcal{J}_{L'_R} = -ie^{ik_3z}\int\limits_{\rho}^{R} \left\{\frac{H_1^{(1)}(e^{i\pi} v_{32}R_0)}{e^{i\pi}v_{32}H_0^{(1)}(e^{i\pi}v_{32}R_0)\,\Phi_1(k_3+i\tau) - H_1^{(1)}(e^{i\pi}v_{32}R_0)\,\Phi_2(k_3+i\tau)} - \right.$$

$$\left. - \frac{H_1^{(1)}(v_{32}R_0)}{v_{32}H_0^{(1)}(v_{32}R_0)\,\Phi_1(k_3+i\tau) - H_1^{(1)}(v_{32}R_0)\,\Phi_2(k_3+i\tau)}\right\} e^{-z\tau} d\tau. \tag{35}$$

Using the formula

$$H_p^{(1)}(e^{i\pi}z) = -e^{-i\pi p} H_p^{(2)}(z)$$

and letting $\rho \to 0$ and $R \to 0$, we obtain

$$\mathcal{J}_{L'_R} = -ie^{ik_3z}\int\limits_0^{\infty} \left[\frac{H_1^{(2)}(v_{32}R_0)}{v_{32}H_0^{(2)}(v_{32}R_0)\,\Phi_1(k_3+i\tau) - H_1^{(2)}(v_{32}R_0)\,\Phi_2(k_3+i\tau)} - \right.$$

$$\left. - \frac{H_1^{(1)}(v_{32}R_0)}{v_{32}H_0^{(1)}(v_{32}R_0)\,\Phi_1(k_3+i\tau) - H_1^{(1)}(v_{32}R_0)\,\Phi_2(k_3+i\tau)}\right] e^{-z\tau} d\tau. \tag{36}$$

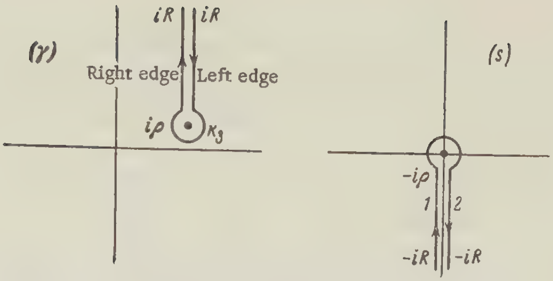


Figure 3.

For small values of the radius R_0, in the frequency range from $6 \cdot 10^3$ to $3 \cdot 10^7$ cps for sea water, moist earth and dry earth, we can replace the cylindrical functions by the first terms in their series expansions. An investigation of the integrand has shown that the upper limit of integration can be adequately confined to the number m, determined from the inequality

$$|v_{32}| R_0 = \left|\sqrt{-i\tau(2k_3 + i\tau)}\right| R_0 \ll 1 \tag{37}$$

or

$$\left|\sqrt{-i\tau(2k_3 + i\tau)}\right| = \frac{1}{\alpha R_0} \quad \text{for} \quad \alpha \gg 1. \tag{38}$$

Assuming that

$$k_3 = \frac{2\pi}{\lambda}\sqrt{\varepsilon'_{k3}} = n + ip, \tag{39}$$

we can rewrite (38) in the form

$$\tau^4 + 4p\tau^3 + 4(n^2 + p^2)\tau^2 - \frac{1}{(\alpha R_0)^4} = 0. \tag{40}$$

The upper limit of the integral, m, is the root of Equation (40). The quantity $\tau = m$ is measured in nepers/meter.

Table 2

Frequency cps	Sea Water $\left(\varepsilon'_3 = 80,\ \sigma_3 = 4\,\frac{1}{\text{ohm}\cdot\text{m}}\right)$	Moist Earth $\left(\varepsilon'_3 = 10,\ \sigma_3 = 5\cdot 10^{-2}\,\frac{1}{\text{ohm}\cdot\text{m}}\right)$	Dry Earth $\left(\varepsilon'_3 = 4,\ \sigma_3 = 10^{-3}\,\frac{1}{\text{ohm}\cdot\text{m}}\right)$
$6\cdot 10^3$	16,35601	16.65725	16,65127
$3\cdot 10^4$	15.97217	16.59444	16.63224
$3\cdot 10^5$	14,44514	16.42194	16.55784
$3\cdot 10^6$	10.08883	15.89038	16.33311
$3\cdot 10^7$	4.29713	14.13788	15.75743

Table 2 lists the roots of Equation (40), calculated for sea water, moist earth, and dry earth with $R_0 = 6$ mm and $\alpha = 10$.

After carrying out the transformations and discarding terms of higher order of smallness, we have

$$
\begin{aligned}
\mathcal{J}_{L'_R} = -i\frac{\pi}{2}e^{ik_3z}\Bigg\{ & \int\limits_0^m \frac{e^{-z\tau}d\tau}{\frac{\pi}{2}\tau(2k_3+i\tau)-\frac{\tau}{2}i(2k_3+i\tau)\ln[-i\tau(2k_3+i\tau)R_0^2]+} \rightarrow \\
& \rightarrow \frac{e^{-z\tau}d\tau}{+(k_3^2-k_2^2+i2k_3\tau-\tau^2)\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0}} - \\
& -\int\limits_0^m \frac{e^{-z\tau}d\tau}{-\frac{\pi}{2}\tau(2k_3+i\tau)-i\frac{\tau}{2}(2k_3+i\tau)\ln[-i\tau(2k_3+i\tau)R_0^2]+} \rightarrow \\
& \rightarrow \frac{e^{-z\tau}d\tau}{+(k_3^2-k_2^2+i2k_3\tau-\tau^2)\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0}} \Bigg\}.
\end{aligned}
\tag{41}
$$

This integral can be calculated approximately through each medium separately.

a) Sea Water. In the frequency range from $6 \cdot 10^3$ to $3 \cdot 10^7$ cps, we can neglect in the denominator of the integrand of (41) the term $-i\tau(2k_3+i\tau)\ln[-i\tau(2k_3+i\tau)R_0^2]$. As a result, we obtain

$$
\mathcal{J}_{L'_R} = -i\frac{\pi}{2}e^{ik_3z}\left\{\int\limits_0^m \frac{e^{-z\tau}d\tau}{a_1\tau^2+b_1\tau+c_1} - \int\limits_0^m \frac{e^{-z\tau}d\tau}{a_2\tau^2+b_2\tau+c_1}\right\},
\tag{42}
$$

where

$$
a_1 = i\frac{\pi}{2} - \frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0};
\tag{43}
$$

$$
b_1 = k_3\left(\pi + i2\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0}\right);
\tag{44}
$$

$$
a_2 = -i\frac{\pi}{2} - \frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0};
\tag{45}
$$

$$
b_2 = k_3\left(-\pi + i2\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0}\right);
\tag{46}
$$

$$
c_1 = (k_3^2-k_2^2)\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0}.
\tag{47}
$$

We have the function

$$
\frac{1}{a\tau^2+b\tau+c} = \sum_{n=0}^{\infty} A_n\tau^n.
\tag{48}
$$

Thus, we obtain

$$
\mathcal{J}_{L'_R} = -i\frac{\pi}{2}e^{ik_3z}\left(\int\limits_0^m \sum_{n=0}^{\infty} A_n^{(1)}\tau^n e^{-z\tau}d\tau - \int\limits_0^m \sum_{n=0}^{\infty} A_n^{(2)}\tau^n e^{-z\tau}d\tau\right).
\tag{49}
$$

After calculating the coefficients of the series (49), we have

$$
\mathcal{J}_{L'_R} = -i\frac{\pi}{2}e^{ik_3z}\left\{\frac{b_2-b_1}{c_1^2}\int\limits_0^m \tau e^{-z\tau}d\tau + \frac{a_2c_1-b_2^2-a_1c_1+b_1^2}{c_1^3}\int\limits_0^m \tau^2e^{-z\tau}d\tau + \ldots\right\}.
\tag{50}
$$

Restricting the series to two terms and calculating the integrals in (50), we get

$$\mathcal{J}_{L'_R} = -i\frac{\pi}{2}e^{ik_3 z}\left\{-\frac{b_2-b_1}{c_1^2}\left(\frac{m}{z}+\frac{1}{z^2}\right)e^{-mz} - \frac{(a_2-a_1)c_1+b_1^2-b_2^2}{c_1^3}\times\right.$$
$$\left.\times\left(\frac{m^2}{z}+\frac{2m}{z^2}+\frac{2}{z^3}\right)e^{-mz}+\frac{b_2-b_1}{c_1^2z^2}+2\frac{(a_2-a_1)c_1+b_1^2-b_2^2}{c_1^3z^3}\right\}. \qquad (51)$$

b) Moist earth. In the frequency range from $6\cdot 10^3$ to $3\cdot 10^6$ cps, we can neglect the term $i\tau(2k_3+i\tau)\ln[-i\tau(2k_3+i\tau)R_0^2]$ in the denominator of the integrand of (41).

Consequently, the integral $\mathcal{J}_{L'_R}$ will be determined in this range by Equation (51). In the frequency from $3\cdot 10^6$ to $3\cdot 10^7$ cps, Equation (41) is best rewritten

$$\mathcal{J}_{L'_R} = -i\frac{\pi}{2}e^{ik_3 z}(\mathcal{J}_1+\mathcal{J}_2), \qquad (52)$$

where

$$\mathcal{J}_1 = \int_0^1 \frac{e^{-z\tau}d\tau}{\frac{\pi}{2}\tau(2k_3+i\tau)-i\frac{\tau}{2}(2k_3+i\tau)\ln[-i\tau(2k_3+i\tau)R_0^2]+} \rightarrow$$
$$\rightarrow \frac{e^{-z\tau}d\tau}{+(k_3^2-k_2^2+i2k_3\tau-\tau^2)\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0}} -$$
$$-\int_0^1 \frac{e^{-z\tau}d\tau}{-\frac{\pi}{2}\tau(2k_3+i\tau)-i\frac{\tau}{2}(2k_3+i\tau)\ln[-i\tau(2k_3+i\tau)R_0^2]+} \rightarrow$$
$$\rightarrow \frac{e^{-z\tau}d\tau}{+(k_3^2-k_2^2+i2k_3\tau-\tau^2)\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0}}; \qquad (53)$$

$$\mathcal{J}_2 = \int_1^m \frac{e^{-z\tau}d\tau}{\frac{\pi}{2}\tau(2k_3+i\tau)-i\frac{\tau}{2}(2k_3+i\tau)\ln[-i\tau(2k_3+i\tau)R_0^2]+} \rightarrow$$
$$\rightarrow \frac{e^{-z\tau}d\tau}{+(k_3^2-k_2^2+i2k_3\tau-\tau^2)\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0}} -$$
$$-\int_1^m \frac{e^{-z\tau}d\tau}{-\frac{\pi}{2}\tau(2k_3+i\tau)-i\frac{\tau}{2}(2k_3+i\tau)\ln[-i\tau(2k_3+i\tau)\cdot R_0^2]+} \rightarrow$$
$$\rightarrow \frac{e^{-z\tau}d\tau}{+(k_3^2-k_2^2+i2k_3\tau-\tau^2)\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0}}. \qquad (54)$$

Calculations have shown that in the evaluation of the integral we can neglect in the denominator of the integrand the term $i\frac{\tau}{2}(2k_3+i\tau)\cdot\ln[-i\tau(2k_3+i\tau)R_0^2]$.

As a result the value of the integral $\mathcal{J}_1$ is

$$\mathcal{J}_1 = -\frac{b_2-b_1}{c_1^2}\left(\frac{1}{z}+\frac{1}{z^2}\right)e^{-z} - \frac{(a_2-a_1)c_1+b_1^2-b_2^2}{c_1^3}\left(\frac{1}{z}+\frac{2}{z^2}+\frac{2}{z^3}\right)e^{-z}+$$
$$+\frac{b_2-b_1}{c_1^2z^2}+2\frac{(a_2-a_1)c_1+b_1^2-b_2^2}{c_1^3z^3}. \qquad (55)$$

With increasing τ, the modulus of the denominator of the integrand (55) increases. When $\tau > 1$, the value of the modulus of the denominator of the integrand in (55) is large, whereas the modulus of the exponential function in the numerator of the integrand is quite small when $z > 1$.

Consequently, when z > 1 the integral $\mathcal{J}_1$ yields the principal part of the solution, whereas the integral $\mathcal{J}_2$ is only a correction. Obviously, we can neglect the integral $\mathcal{J}_2$ in the calculations when z > 1. The integral $\mathcal{J}_2$ becomes significant only when z < 1.

The integral $\mathcal{J}_2$ must be calculated numerically. However, taking into account the monotonic character of the variation of the logarithmic function as τ increases from 1 to m, we can approximately set $\tau = 1$ under the logarithm sign in the calculation of the integral $\mathcal{J}_2$ i.e.,

$$\mathcal{J}_2 = \int_1^m \frac{e^{-z\tau}d\tau}{\frac{\pi}{2}\tau(2k_3 + i\tau) - i\frac{\tau}{2}(2k_3 + i\tau)\ln[(-i2k_3 + 1)R_0^2] +} \rightarrow$$

$$\rightarrow \frac{e^{-z\tau}d\tau}{+(k_3^2 - k_2^2 + i2k_3\tau - \tau^2)\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0}} -$$

$$-\int_1^m \frac{e^{-z\tau}d\tau}{-\frac{\pi}{2}\tau(2k_3 + i\tau) - i\frac{\tau}{2}(2k_3 + i\tau)\ln[(-i2k_3 + 1)R_0^2] +} \rightarrow$$

$$\rightarrow \frac{e^{-z\tau}d\tau}{+(k_3^2 - k_2^2 + i2k_3\tau - \tau^2)\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0}}. \tag{56}$$

As a result we obtain

$$\mathcal{J}_2 = \int_1^m \frac{e^{-z\tau}d\tau}{a_3\tau^2 + b_3\tau + c_2} - \int_1^m \frac{e^{-z\tau}d\tau}{a_4\tau^2 + b_4\tau + c_2}, \tag{57}$$

where

$$a_3 = i\frac{\pi}{2} + \frac{1}{2}\ln[(-i2k_3 + 1)R_0^2] - \frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0};$$

$$b_3 = \pi k_3 - ik_3\ln[(-i2k_3 + 1)R_0^2] + i2k_3\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0};$$

$$a_4 = -i\frac{\pi}{2} + \frac{1}{2}\ln[(-i2k_3 + 1)R_0^2] - \frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0};$$

$$b_4 = -\pi k_3 - ik_3\ln[(-i2k_3 + 1)R_0^2] + i2k_3\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0};$$

$$c_2 = (k_3^2 - k_2^2)\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0}.$$

In analogy with the preceding, the expression (57) is written as follows

$$\mathcal{J}_2 = \frac{b_4 - b_3}{c_2^2}\int_1^m \tau e^{-z\tau}d\tau + \frac{a_4c_2 - b_4^2 - a_3c_2 + b_3^2}{c_2^3}\int_1^m \tau^2 e^{-z\tau}d\tau + \ldots. \tag{58}$$

We terminate the series with the second term and evaluate the integral and get

$$\mathcal{J}_2 = -\frac{b_4 - b_3}{c_2^2}\left[e^{-mz}\left(\frac{m}{z} + \frac{1}{z^2}\right) - e^{-z}\left(\frac{1}{z} + \frac{1}{z^2}\right)\right] -$$

$$-\frac{a_4c_2 - b_4^2 - a_3c_2 + b_3^2}{c_2^3}\left[e^{-mz}\left(\frac{m^2}{z} + \frac{2m}{z^2} + \frac{2}{z^3}\right) - e^{-z}\left(\frac{1}{z} + \frac{2}{z^2} + \frac{2}{z^3}\right)\right]. \tag{59}$$

The distance z in (55) and (59) is in meters.

c) Dry Earth. In the frequency range up to 3 $\cdot 10^5$ cps inclusive, the integral (41) has the same form as (52). The integral $\mathcal{J}_1$ can be calculated approximately from (55), while $\mathcal{J}_2$ is approximately given by (59).

At frequencies higher than 3 . 10^5 cps, the integral $\mathcal{J}_1$ must be evaluated numerically.

Calculation of the integral along the cut L''$_R$. The integral over the cut L''$_R$ is written in the form

$$\mathcal{J}_{L''_R} = \int\limits_{L''_R} \frac{H_1^{(1)}(v_3R_0)\, e^{i\gamma z} d\gamma}{v_2 [\mathscr{J}_0(v_2r_0) N_1(v_2R_0) - \mathscr{J}_1(v_2R_0) N_0(v_2r_0)]\, v_3H_0^{(1)}(v_3R_0) -} \rightarrow$$

$$\rightarrow \frac{H_1^{(1)}(v_3R_0)\, e^{i\gamma z} d\gamma}{- v_2^2 [\mathscr{J}_0(v_2r_0) N_0(v_2R_0) - \mathscr{J}_0(v_2R_0) N_0(v_2r_0)] \frac{\varepsilon_{k3}}{\varepsilon_{k2}} H_1^{(1)}(v_3R_0)}. \tag{60}$$

Making the substitution

$$k_2 - \gamma = \sigma, \quad d\gamma = - d\sigma,$$

we obtain

$$\mathcal{J}_{L''_R} = - ie^{ik_2z} \int\limits_{\rho}^{R} \Bigg\{ \frac{H_1^{(1)}(v_3R_0)}{v_{21} [\mathscr{J}_0(v_{21}r_0) N_1(v_{21}R_0) - \mathscr{J}_1(v_{21}R_0) N_0(v_{21}r_0)]\, v_3H_0^{(1)}(v_3R_0) -} \rightarrow$$

$$\rightarrow \frac{H_1^{(1)}(v_3R_0)}{- v_{21}^2 [\mathscr{J}_0(v_{21}r_0) N_0(v_{21}R_0) - \mathscr{J}_0(v_{21}R_0) N_0(v_{21}r_0)] \frac{\varepsilon_{k3}}{\varepsilon_{k2}} H_1^{(1)}(v_3R_0)} -$$

$$- \frac{H_1^{(1)}(v_3R_0)}{v_{22} [\mathscr{J}_0(v_{22}r_0) N_1(v_{22}R_0) - \mathscr{J}_1(v_{22}R_0) N_0(v_{22}r_0)]\, v_3H_0^{(1)}(v_3R_0) -} \rightarrow$$

$$\rightarrow \frac{H_1^{(1)}(v_3R_0)}{- v_{22}^2 [\mathscr{J}_0(v_{22}r_0) N_0(v_{22}R_0) - \mathscr{J}_0(v_{22}R_0) N_0(v_{22}r_0)] \frac{\varepsilon_{k3}}{\varepsilon_{k2}} H_1^{(1)}(v_3R_0)} \Bigg\} e^{-uz}\, du, \tag{61}$$

where

$$v_3 = \sqrt{k_3^2 - (k_2 + iu)^2}; \quad v_{21} = \sqrt{ue^{i\frac{3\pi}{2}}(2k_2 + iu)};$$

$$v_{22} = \sqrt{ue^{-i\frac{\pi}{2}}(2k_2 + iu)}.$$

Using the formula

$$\mathscr{J}_p(e^{i\pi}z) = e^{i\pi p}\mathscr{J}_p(z),$$

$$N_p(e^{i\pi}z) = e^{-i\pi p}N_p(z) + 2i \cos p\pi\, \mathscr{J}_p(z)$$

and letting $\rho \to 0$ and R $\to \infty$, we obtain

$$\mathcal{J}_{L''_R} = - ie^{ik_2z} \int\limits_{0}^{\infty} \Bigg\{ \frac{H_1^{(1)}(v_3R_0)}{v_{22} [\mathscr{J}_0(v_{22}r_0) N_1(v_{22}R_0) - \mathscr{J}_1(v_{22}R_0) N_0(v_{22}r_0)]\, v_3H_0^{(1)}(v_3R_0) -} \rightarrow$$

$$\rightarrow \frac{H_1^{(1)}(v_3R_0)}{- v_{22}^2 [\mathscr{J}_0(v_{22}r_0) N_0(v_{22}R_0) - \mathscr{J}_0(v_{22}R_0) N_0(v_{22}r_0)] \frac{\varepsilon_{k3}}{\varepsilon_{k2}} H_1^{(1)}(v_3R_0)} -$$

$$- \frac{H_1^{(1)}(v_3R_0)}{v_{22} [\mathscr{J}_0(v_{22}r_0) N_1(v_{22}R_0) - \mathscr{J}_1(v_{22}R_0) N_0(v_{22}r_0)]\, v_3H_0^{(1)}(v_3R_0) -} \rightarrow$$

$$\rightarrow \frac{H_1^{(1)}(v_3R_0)}{- v_{22}^2 [\mathscr{J}_0(v_{22}r_0) N_0(v_{22}R_0) - \mathscr{J}_0(v_{22}R_0) N_0(v_{22}r_0)] \frac{\varepsilon_{k3}}{\varepsilon_{k2}} H_1^{(1)}(v_3R_0)} \Bigg\} e^{-uz} du = 0. \tag{62}$$

Determination of the residue of the integrand. We set

$$\varphi(\gamma) = \frac{v_3}{v_2^3} H_1^{(1)}(v_3R_0)\, e^{i\gamma z}. \tag{63}$$

The residue of the integrand at the point $\gamma = \gamma_{p1}$ is

$$R(z) = \frac{\varphi(\gamma)\big|_{\gamma=\gamma_{p1}}}{\dfrac{dD}{d\gamma}\Big|_{\gamma=\gamma_{p1}}}. \tag{64}$$

Let us determine $\dfrac{dD}{d\gamma}\Big|_{\gamma=\gamma_{p1}}$. In the expression (19) we replace the cylindrical functions by the first terms of their series expansions. As a result we obtain

$$D = \frac{v_3^2}{v_2^2}\left(1 - i\frac{2}{\pi}\ln\frac{1{,}12}{v_3R_0}\right)\left[-\frac{2}{\pi}\frac{1}{v_2R_0} + \frac{v_2R_0}{2}\frac{2}{\pi}\ln\frac{1{,}12}{v_2r_0}\right] - \frac{\varepsilon_{k3}}{\varepsilon_{k2}}\frac{v_3}{v_2}\left(\frac{v_3R_0}{2} - i\frac{2}{\pi}\frac{1}{v_3R_0}\right)\left[-\frac{2}{\pi}\ln\frac{1{,}12}{v_2R_0} + \frac{2}{\pi}\ln\frac{1{,}12}{v_2r_0}\right]. \tag{65}$$

Discarding the higher-order terms $v_2R_0/2 \cdot 2/\pi$ In $1.12/\, v_0r_0$ and $v_3R_0/2$, we have

$$D = i\frac{4}{\pi^2R_0}\frac{\varepsilon_{k3}}{\varepsilon_{k2}v_2}\left\{\frac{v_3^2}{v_2^2}\frac{\varepsilon_{k2}}{\varepsilon_{k3}}\left(\ln\frac{1{,}12}{v_3R_0} + i\frac{\pi}{2}\right) + \ln\frac{R_0}{r_0}\right\}. \tag{66}$$

Differentiating (66) and recognizing that

$$\ln\frac{1{,}12}{v_3R_0} + i\frac{\pi}{2} = -\frac{\varepsilon'_{k3}}{\varepsilon'_{k2}}\frac{k_0^2\varepsilon'_{k2} - \gamma^2}{k_0^2\varepsilon'_{k3} - \gamma^2}\ln\frac{R_0}{r_0}, \tag{67}$$

we obtain

$$\frac{dD}{d\gamma}\Big|_{\gamma=\gamma_{p1}} = i\frac{4\gamma}{\pi^2R_0v_2^3}\left(-2\frac{v_3^2 - v_2^2}{v_3^2}\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0} + 1\right)\Big|_{\gamma=\gamma_{p1}}. \tag{68}$$

Substituting (68) in (64) and dropping the index of γ_{p1}, we obtain

$$R(z) = -i\frac{\pi^2R_0v_3H_1^{(1)}(v_3R_0)\,e^{i\gamma z}}{4\gamma\left(-2\dfrac{v_3^2 - v_2^2}{v_3^2}\dfrac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\dfrac{R_0}{r_0} + 1\right)}. \tag{69}$$

Assuming

$$R(z) \simeq -\frac{\pi e^{i\gamma z}}{2\gamma\left(-2\dfrac{v_3^2 - v_2^2}{v_3^2}\dfrac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\dfrac{R_0}{r_0} + 1\right)}. \tag{70}$$

we get

$$H_1^{(1)}(v_3R_0) \simeq -i\frac{2}{\pi}\frac{1}{v_3R_0},$$

Thus, the integral (31) has a value

$$\mathcal{J} = -\frac{i\pi^2e^{i\gamma z}}{\gamma\left(-2\dfrac{v_3^2 - v_2^2}{v_3^2}\dfrac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\dfrac{R_0}{r_0} + 1\right)} + \mathcal{J}_{L'_R}. \tag{71}$$

Substituting (71) in (28) we obtain the following expression for the current in the wire

$$I_z = -i\frac{2\omega\varepsilon_{k3}\mathcal{E}}{\pi}\left(\mathcal{J}_p + \mathcal{J}_{L'_R}\right), \tag{72}$$

where

$$\mathcal{J}_p = -\frac{i\pi^2 e^{i\gamma z}}{\gamma\left(-2\frac{v_3^2-v_2^2}{v_3^2}\frac{\varepsilon_{k3}}{\varepsilon_{k2}}\ln\frac{R_0}{r_0}+1\right)}. \tag{73}$$

3. INVESTIGATION OF THE EQUATION FOR THE CURRENT IN AN INSULATED WIRE

The contour integral of expression (28) for the current in an insulated wire contains a pole and two branch points $\gamma = \pm k_2$ and $\gamma = \pm k_3$.

The part of the solution corresponding to the pole is a current wave, whose propagation along the wire is characterized by a phase factor $e^{i\gamma z}$. The amplitude of this wave decreases exponentially as $e^{-\beta z}$. This indicates that a surface electromagnetic wave propagates along the insulated wire located in the conducting medium. This wave is sometimes called the cable wave.

As is well known, the telegrapher's equations pertain to the propagation of the cable wave along the line. The first term $\mathcal{J}_p$ of expression (72) corresponding to the residue, satisfies the telegraphy equations.

Part of the solution $\mathcal{J}_{L'_R}$, corresponding to the singular point $\gamma = \pm k_3$ in the contour integral is a space wave. The amplitude of this wave decreases with z in a rather complicated fashion.

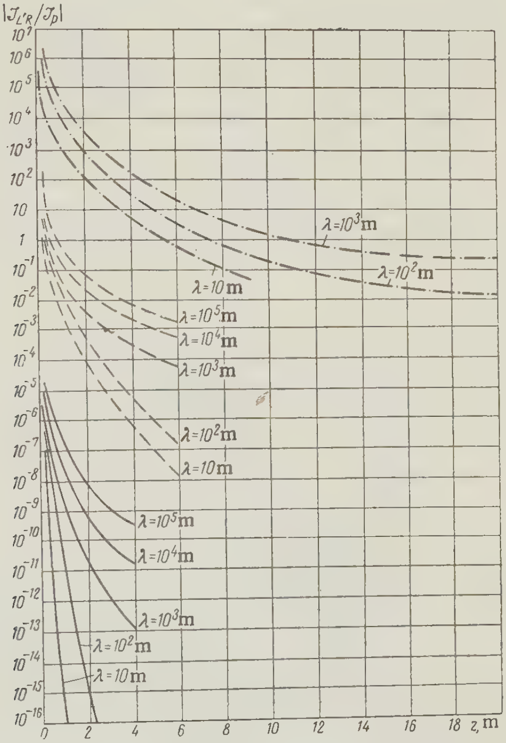


Figure 4.

In order to establish which of the current waves predominates in an insulating wire situated in a conducting medium, we have investigated the modulus of the ratio $|\mathcal{J}_{L'_R}/\mathcal{J}_p|$ at short, medium, and long wavelengths for sea water, moist earth, and dry earth. The results of the calculation are shown in Figure 4. (solid curves — sea water, dashes — moist earth; dash-dots — dry earth).

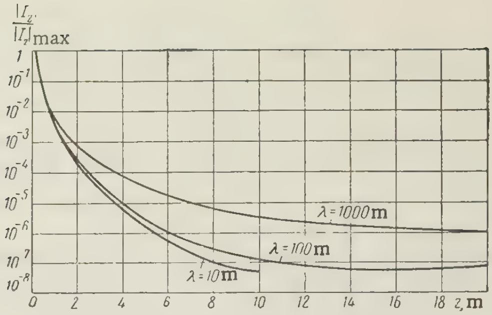


Figure 5

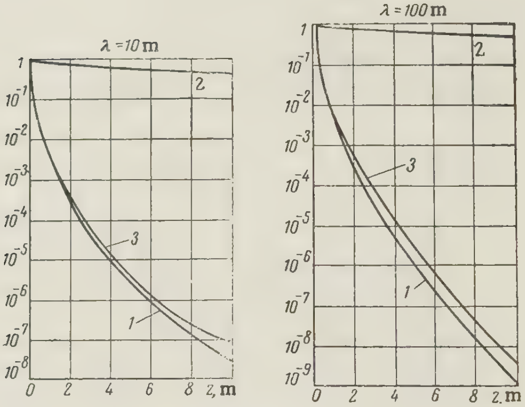


Figure 6

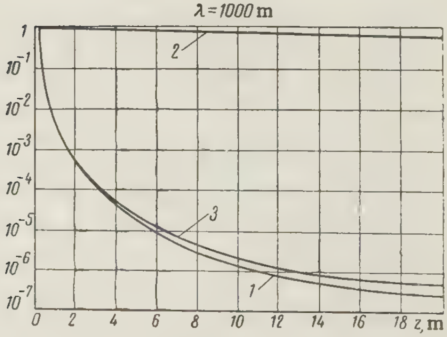


Figure 7

It follows from the figures that when the insulated wire is placed in sea water or in moist or dry air, the value of $|\mathcal{J}_{L'_R}/\mathcal{J}_p|$ decreases with increasing frequency for a fixed distance from the point where the voltage is applied.

For an insulated wire situated in sea water and in moist air, we have $|\mathcal{J}_{L'_R}/\mathcal{J}_p| > 1$ at frequencies from $6 \cdot 10^3$ to $3 \cdot 10^7$ cps only at very small distances from the voltage source $(z < 1)$. Consequently, with exception of a small initial portion of the wire, the cable wave predominates over the space wave. On this basis we can conclude that the telegrapher's equations are applicable to the theory of an insulated wire situated in sea water and in moist earth.

For an insulated wire located in dry earth, the inequality $|\mathcal{J}_{L'_R}/\mathcal{J}_p| < 1$ sets in at frequencies below $3 \cdot 10^7$ at a relatively short distance from the point of application of the voltage. The lower the frequency, the greater this distance.

The distribution of the current amplitude along an insulated wire in dry earth, shown in Figure 5, differ from the exponential curves.

Normalized plots of the absolute values of the functions $1 - |\mathcal{J}_{L'_R}|$, $2 - |e^{i\varkappa_3 z}|$, $3 - |\varphi(z)|$ (Figures 6 and 7), of the expression (52) for the current $\mathcal{J}_{L'_R}$, for an insulated wire situated in dry earth, show that in the initial portion of the wire the current $\mathcal{J}_{L'_R}$ is characterized essentially by a rapidly varying function $\varphi(z)$ in the frequency range $3 \cdot 10^5 - 3 \cdot 10^7$ cps.

The author is indebted to L.A. Vaynshteyn for valuable advice and a discussion of the results.

REFERENCES

1. V.N. Kessenikh, Dokl. AN SSSR, 1940, 27, 6, 558.
2. V.V. Vladimirskiy, Izv. AN SSSR, ser. fiz (News of the Academy of Sciences USSR, Physics Series), 1944, 8, 3, 139.
3. L.S. Korol'kevich, G.A. Lavrov, Tr. TsNIIS SA (Transactions, Central Scientific Research Institute for Communication), 1951, 2, (85).

Submitted to the editors 16 July 1960

TRANSIENT CHARACTERISTICS OF COAXIAL SYSTEM WITH PERIODIC STRUCTURE

S. I. Viglin

The telegrapher's equations are set up for the equivalent circuit of a magnetodielectric line, the form of the transfer characteristic is evaluated, and the duration of the signal buildup at any point of the line is determined.

INTRODUCTION

Many constructions of coaxial systems with periodic structure, suitable for shaping and delaying short video pulses, have been proposed recently [1, 2, 3]. Unlike the homogeneous coaxial line, these constructions have a greater running inductance for the same radial dimensions. Figures 1 and 2 show two modifications of such systems.

In the helical line (Figure 1), the increase in the running inductance is attained by replacing the straight central conductor by a solenoid. In some constructions of this type, the outer cylinder is made in the form of a solenoid.

In a magnetodielectric line (Figure 2), the running inductance is increased by filling the space between the round central conductor 5 and the external metallic cylinder 6 with a set of metallic washers 2, made of a material having a high permeability. These wahsers are insulated from each other by dielectric liners 3, and are insulated from the internal rod 5 and from the external cylinder 6 by dielectric bushings 1 and 4. A detailed investigation of the propagation of electromagnetic waves in a magnetodielectric line and of its properties as a shaping and delay device is given in reference [4].

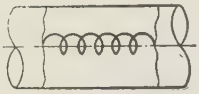


Figure 1. Helical line.

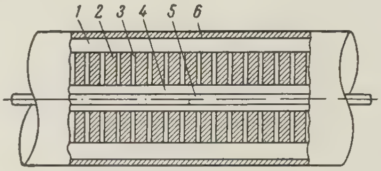


Figure 2. Magnetodielectric line.

Owing to the increase in the running inductance in the periodic coaxial systems, the speed of the propagation of the electromagnetic waves v is decreased and the wave resistance ρ is increased, an important factor in the design of shaping and delay devices [5, 6].

References [2, 4] give approximate equations for the running parameters of these systems. In the design, however, it is also necessary to estimate the distortion of the pulses, and in particular to determine the duration of the front of the load pulses. This problem is solved in the present article.

An investigation of the distortions of the transmitted signals is complicated by the fact that the coaxial systems under consideration are characterized by strong dispersion, this being due principally to the presence of a longitudinal capacitance C_t, and to the presence of active losses.

In a magnetodielectric line the longitudinal capacitance is brought about by the longitudinal electric field between the neighboring metallic washers 2, and in a helical line the cause is the field between neighboring turns of the solenoid. Since this capacitance couples the different line elements, an electric field is produced simultaneously in several neighboring line elements by the wave propagating through the line, and this distorts the front of the wave and consequently the form of the transmitted signal. The dispersion is caused by the fact that the capacitive reactance $1/2\pi C_t$ (f is the frequency of the transmitted signal) is a function of the frequency.

The effect of active losses on the change in the phase velocity is appreciable only at very high frequencies (above 100 Mc in the case of a magnetodielectric line [4]). In the video range we can, therefore, confine ourselves to an investigation of the dependence of the distortion on the longitudinal capacitance C_t only.

1. DIFFERENTIAL EQUATIONS

An approximate equivalent circuit of a magnetodielectric line [4] is shown in Figure 3. Here L, C. and C_t are respectively the running inductance, capacitance, and longitudinal capacitance of the line; Δx is the period of the structure (length of the line element along the wave propagation axis, x), u and i are the instantaneous values of the voltage and current in the line.

The complete equivalent circuit should also include the inductance L_d due to the magnetic field in the dielectric. This inductance is connected in series with the resonant circuit ($L\Delta x$, $C_t\Delta x$). Inasmuch as an account of L_d would increase the order of the differential equations and would make the solution more complicated, we disregard it in the present paper. This neglect is justified, since $L_d \ll L$, because the permeability of the metallic washers is $\mu \gg 1$.

The equivalent circuit of the helical line is of similar form. We note that many problems in the theory of transients in transformer windings can be reduced to an investigation of a similar circuit [7].

Kirchhoff's equations for the circuit shown in Figure 3 have the form

$$\Delta u = L\Delta x \frac{\partial i_L}{\partial t}, \qquad \Delta i = i_C = C\Delta x \frac{\partial u}{\partial t}. \tag{1}$$

Since the period of the structure Δx is always a sufficiently small quantity ($\Delta x \ll \lambda$, where λ is the wavelength), we can put

$$\Delta u = \frac{\partial u}{\partial x}\Delta x, \qquad \Delta i = \frac{\partial i}{\partial x}\Delta x.$$

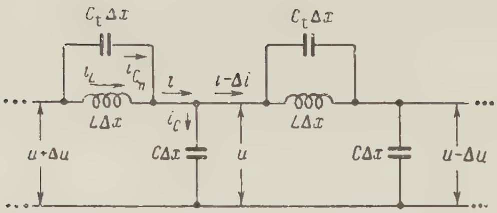


Figure 3. Equivalence diagram of a magnetodielectric line

$$\frac{\partial u}{\partial x} = L\frac{\partial i_L}{\partial t}, \tag{2}$$

$$\frac{\partial i}{\partial x} = C\frac{\partial u}{\partial t}. \tag{3}$$

The current i_L is given by

$$i_L = i - i_{C_t} = i - C_t \Delta x \frac{\partial (\Delta u)}{\partial t} = i - C_t \Delta x^2 \frac{\partial^2 u}{\partial x\, \partial t}.$$

Substituting this expression in (2), we get

$$\frac{\partial u}{\partial x} = L\frac{\partial i}{\partial t} - \frac{1}{\omega_{\text{lim}}^2}\frac{\partial^3 u}{\partial x\, \partial t^2}, \tag{4}$$

where

$$\omega_{\text{lim}} = \frac{1}{\Delta x \sqrt{LC_t}} \tag{5}$$

is the resonant frequency of the circuit ($L\Delta x$, $C_t \Delta x$).

Equations (3) and (4) are the telegrapher's equations for the circuits shown in Figure 3, with longitudinal capacitance. As $C_t \to 0$ ($\omega_{\text{lim}} \to \infty$), Equation (4) assumes the well-known form for a uniform lossless line.

Let us take the Laplace transform of the system (3) - (4)

$$\bar{f}(p) = p\int_0^\infty f(t)\, e^{-pt}\, dt.$$

We then obtain the following system of equations for the transform $\bar{f}(p)$:

$$\frac{\partial \bar{i}(p)}{\partial x} = pC\bar{u}(p), \tag{6}$$

$$\frac{\partial \bar{u}(p)}{\partial x} = pL\bar{i}(p) - \frac{p^2}{\omega_{\text{lim}}^2}\frac{\partial \bar{u}(p)}{\partial x}. \tag{7}$$

Differentiating both halves of (7) with respect to x and substituting $\partial \bar{i}(p)/\partial x$ into (6) we get

$$\frac{\partial^2 \bar{u}(p)}{\partial x^2} = \gamma^2(p)\, \bar{u}(p), \tag{8}$$

where

$$\gamma(p) = \frac{p\sqrt{LC}}{\sqrt{1 + \frac{p^2}{\omega_{\text{lim}}^2}}}.$$

We introduce a new variable

$$q = \frac{p}{\omega_{\text{lim}}}, \tag{9}$$

corresponding to the dimensionless time

$$\tau = \omega_{\text{lim}} t. \tag{10}$$

Then

$$\gamma(q) = \frac{\omega_{\text{lim}} q \sqrt{LC}}{\sqrt{1+q^2}}.$$

2. TRANSIENT CHARACTERISTIC

We examine the solution of (8) for an infinitely long line, along which a traveling wave propagates. Then

$$\overline{u}\,(x,\,q) = Ae^{-\gamma(q)x}.$$

When x = 0 we have $\overline{u}(0,q) = A$. In order to find the transient characteristic $h(x,\tau)$, we assum that the input signal is a step function

$$u\,(0,\,\tau) = 1\,(\tau).$$

Then

$$A = 1$$

This means that the transform of the transient characteristic is

$$K\,(x,\,q) = e^{-\gamma(q)x}.$$

Putting

$$\alpha = \omega_{\lim} \sqrt{LC}\,x, \tag{11}$$

we obtain

$$K\,(\alpha,\,q) = e^{-\frac{\alpha q}{\sqrt{1+q^2}}}. \tag{12}$$

In order to find the inverse transform of (12), we apply the Efros transformation [8].

$$F\,[s\,(q)] \doteqdot s\,(q) \int\limits_0^\infty \psi\,(\tau,\,\xi)\,\varphi\,(\alpha,\,\xi)\,d\xi, \tag{13}$$

where

$$\varphi\,(\alpha,\,\xi) \doteqdot F\,(\alpha,\,s); \quad \psi\,(\tau,\,\xi) \doteqdot e^{-\xi s(q)}. \tag{14}$$

The symbol ≑ denotes correspondence between the original and the transform

We set

$$s = s\,(q) = \frac{1+q^2}{q^2}.$$

Then

$$K\,(\alpha,\,s) = e^{-\frac{\alpha}{\sqrt{s}}}.$$

Assuming that

$$F\,(\alpha,\,s) = \sqrt{s}K\,(\alpha,\,s) = \sqrt{s}e^{-\frac{\alpha}{\sqrt{s}}}, \tag{15}$$

we obtain on the basis of the Efros formula (13)

$$\sqrt{s}\,K\,(\alpha,\,s) \doteqdot s \int\limits_0^\infty \psi\,(\tau,\,\xi)\,\varphi\,(\alpha,\,\xi)\,d\xi$$

or

$$\frac{K\,(\alpha,\,s)}{\sqrt{s}} = \frac{e^{-\frac{\alpha}{\sqrt{s}}}}{\sqrt{s}} \doteqdot \int\limits_0^\infty \psi\,(\tau,\,\xi)\,\varphi\,(\alpha,\,\xi)\,d\xi.$$

Let us integrate the resultant expression with respect to the parameter α. This can be done on the basis of the well-known theorems of operational calculus [8]:

$$\int\limits_0^\alpha \frac{e^{-\frac{\alpha}{\sqrt{s}}}}{\sqrt{s}}\,d\alpha \doteqdot \int\limits_0^\alpha \int\limits_0^\infty \psi\,(\tau,\,\xi)\,\varphi\,(\alpha,\,\xi)\,d\xi d\alpha.$$

Interchanging the order of integration in the right side, and recognizing that only the function $\varphi(\alpha,\xi)$ depends on α, we obtain

$$1 - e^{-\frac{\alpha}{\sqrt{s}}} \div \int_0^\infty \psi(\tau, \xi)\, \varphi_1(\alpha, \xi)\, d\xi, \tag{16}$$

where

$$\varphi_1(\alpha, \xi) = \int_0^\alpha \varphi(\alpha, \xi)\, d\alpha.$$

From formula (16) we readily obtain

$$e^{-\frac{\alpha}{\sqrt{s}}} \div h(\alpha, \tau) = 1 - \int_0^\infty \psi(\tau, \xi)\, \varphi_1(\alpha, \xi)\, d\xi. \tag{17}$$

The functions $\Psi(\tau, \xi)$ and $\varphi(\alpha, \xi)$ are the originals of the transforms defined by relations (14) and (15) :

$$\psi(\tau, \xi) \div e^{-\xi s(q)} = e^{-\xi\left(1+\frac{1}{q^2}\right)} = e^{-\xi} e^{-\frac{\xi}{q^2}},$$

$$\varphi(\alpha, \xi) \div F(\alpha, s) = \sqrt{s}\, e^{-\frac{\alpha}{\sqrt{s}}}.$$

Using the handbook [8], we get

$$\varphi(\tau, \xi) = e^{-\xi} \sum_{k=0}^{\infty} \frac{(-1)^k (\tau\sqrt{\xi})^{2k}}{\Gamma(k+1)\,\Gamma(2k+1)},$$

$$\varphi(\alpha, \xi) = \int_0^\infty \frac{x}{2\sqrt{\pi\xi^3}}\, e^{-\frac{x^2}{4\xi}} J_0(2\sqrt{\alpha x})\, dx.$$

Then

$$\varphi_1(\alpha, \xi) = \frac{1}{2\xi}\sqrt{\frac{\alpha}{\pi\xi}} \int_0^\infty \sqrt{x}\, e^{-\frac{x^2}{4\xi}} J_1(2\sqrt{\alpha x})\, dx.$$

If we introduce the variable $y = \alpha x$, we get

$$\varphi_1(\alpha, \xi) = \frac{1}{2\alpha\xi\sqrt{\pi\xi}} \int_0^\infty \sqrt{y}\, e^{-\frac{y^2}{4\alpha^2\xi}} J_1(2\sqrt{y})\, dy.$$

In the calculation of the transfer characteristic $h(\alpha, \tau)$, it is convenient to introduce a new variable $z = \sqrt{\xi}$, and also the transformed functions

$$\Psi(\tau z) = -e^{z^2}\psi(\tau, z) + 1 = -\sum_{k=1}^{\infty} \frac{(-1)^k (\tau z)^{2k}}{\Gamma(k+1)\,\Gamma(2k+1)}, \tag{18}$$

$$\Phi(\alpha z) = \frac{2z}{\alpha}\varphi_1(\alpha, z) = \frac{1}{\alpha^2 z^2\sqrt{\pi}} \int_0^\infty \sqrt{y}\, e^{-\frac{y^2}{4\alpha^2 z^2}} J_1(2\sqrt{y})\, dy. \tag{19}$$

In this case formula (17) assumes the following form, if we recall that $d\xi = 2z\,dz$:

$$h(\alpha, \tau) = 1 - \int_0^\infty e^{-z^2}[1 - \Psi(\tau z)]\, \alpha\Phi(\alpha z)\, dz$$

or

$$h(\alpha, \tau) = 1 - \alpha\int_0^\infty e^{-z^2}\Phi(\alpha z)\, dz + \alpha\int_0^\infty e^{-z^2}\Psi(\tau z)\,\Phi(\alpha z)\, dz. \tag{20}$$

Inasmuch as we have from (18) that $\Psi(\tau, z) = 0$ when $\tau = 0$, the first two terms of (20) yield the initial distribution. Putting

$$h(\alpha, 0) = 1 - \alpha \int_0^\infty e^{-z^2} \Phi(\alpha z)\, dz, \tag{21}$$

we obtain

$$h(\alpha, \tau) = h(\alpha, 0) + \alpha \int_0^\infty e^{-z^2} \Psi(\tau z)\, \Phi(\alpha z)\, dz. \tag{22}$$

This is the final formula and enables us to calculate the transfer characteristic for all values of α and τ.

The functions $\Psi(\tau z)$ and $\Phi(\alpha z)$ are calculated from formulas (18) and (19). The convenience afforded by introducing the transformed functions lies in the fact that each depends on one variable, τz or αz. Parts of these functions are shown in Figure 4.

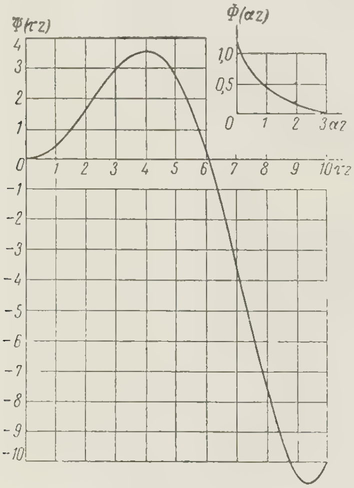


Figure 4. Plots of the transformed functions.

To facilitate the calculations, we can represent the function in series form (see appendix)

$$\Phi(\alpha z) = \sum_{m=0}^{\infty} \frac{(-1)^m (\alpha z)^m}{\Gamma(m+1)\, \Gamma\left(\frac{m+3}{2}\right)}. \tag{23}$$

The quantity $h(\alpha, 0)$ can be determined from (21) letting $q \to \infty$.

$$h(\alpha, 0) = e^{-\alpha}. \tag{24}$$

To find $h(\alpha, \tau)$ as $\tau \to \infty$, we put $q = 0$ in (12). Then

$$h(\alpha, \infty) = 1. \tag{25}$$

We used these formulas to calculate the transfer characteristic $h(\alpha, \tau)$ for values of the parameter $\alpha = 1, 2, 3$, and 5 (Figure 5). The curves plotted enabled us to calculate the active duration of the front τ_f. It was calculated between the levels $h_1(\alpha)$ and $h_2(\alpha)$, defined by the relations

$$\begin{aligned} h_1(\alpha) &= h(\alpha, 0) + 0.1\,[1 - h(\alpha, 0)], \\ h_2(\alpha) &= 1 - 0.1\,[1 - h(\alpha, 0)]. \end{aligned} \tag{26}$$

For sufficiently large α we have $h(\alpha, 0) \simeq 0$; this means that τ_f corresponds to the time interval during which the transfer characteristic rises from $h_1(\alpha) = 0.1$ to $h_2(\alpha) = 0.9$. The funct $\tau_f(\alpha)$ is plotted in Figure 6.

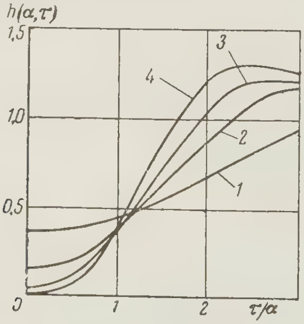


Figure 5. Transfer characteristic: 1 - $\alpha = 1$; 2 - $\alpha = 2$; 3 - $\alpha = 3$, 4 - $\alpha = 5$

Formula (24) shows that at the initial instant $t = 0$ a voltage drop is produced at any

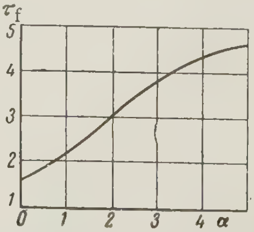


Figure 6. Dependence of active duration of front on the parameter α.

point of the line, corresponding to an instantaneous transmission of the signal. Naturally, this result disagrees with the actual picture of the propagation of an electromagnetic wave along a line. It is brought about by the neglect of the inductance L_d in the equivalent circuit. Nonetheless, the approximate solution enables us to find a correct value of the propagation speed, by measuring the delay time at a level close to 0.5 (this is precisely the level customarily employed in practice).

An analysis of the curves of Figure 5 shows that when $\alpha \geqslant 2$ the curves $h(\alpha, \tau)$ intersect approximately at the point ($\tau/\alpha = 1.04$; $h(\alpha, \tau) = 0.4$). This means that if we measure the time of delay up to the instant when $h(\alpha, \tau)$ reaches a value 0.4, we get

$$\tau_{d_{0.4}} = 1.04\,\alpha \simeq \alpha. \tag{27}$$

Thus, an account of the effect of the capacitance C_t, which leads to the initial jump $h(\alpha, 0)$ at any point of the line, hardly influences the delay time τ_d. Taking (10) and (11) into account, we obtain the real delay time

$$t_d = x_d \sqrt{LC}, \tag{28}$$

where x_d is the length of the delay line. The rate of propagation of the pulse is

$$v = \frac{1}{\sqrt{LC}}. \tag{29}$$

As shown in [4], this velocity corresponds to the phase velocity of propagation of oscillation with frequency f → 0.

The front duration τ_f depends on the parameter α Figure 6 . With increasing x and hence α, the front becomes more gently sloping as the wave propagates. However, as $\alpha \to \infty$ the duration of the front tends to a certain limiting value. In the calculations we can assume

$$\tau_f = 5$$

or

$$t_f = \frac{5}{\omega_{\lim}} = \frac{5}{2\pi f_{\lim}} = \frac{0{,}8}{f_{\lim}}. \tag{30}$$

APPENDIX

Inasmuch as

$$J_1(2\sqrt{y}) = \sum_{m=0}^{\infty} \frac{(-1)^m y^{\frac{2m+1}{2}}}{\Gamma(m+1)\,\Gamma(m+2)},$$

we obtain

$$\Phi(\alpha z) = \frac{1}{\alpha^2 z^2 \sqrt{\pi}} \int_0^{\infty} e^{-\frac{y^2}{4\alpha^2 z^2}} \sum_{m=0}^{\infty} \frac{(-1)^m y^{m+1}}{\Gamma(m+1)\,\Gamma(m+2)}\, dy.$$

Since the series for the Bessel function is absolutely convergent, we can interchange the order of summation and integration. Then

$$\Phi(\alpha, z) = \frac{1}{\alpha^2 z^2 \sqrt{\pi}} \sum_{m=0}^{\infty} \frac{(-1)^m}{\Gamma(m+1)\,\Gamma(m+2)} \int_0^{\infty} y^{m+1} e^{-\frac{y^2}{4\alpha^2 z^2}}\, dy.$$

We introduce the variable

$$x = \frac{y^2}{4\alpha^2 z^2};$$

for which

$$dx = \frac{y\,dy}{2\alpha^2 z^2}, \quad dy = \frac{\alpha z}{\sqrt{x}}\, dx.$$

Substituting the values of y and dy, we rewrite the expression for $\Phi(\alpha z)$ as follows

$$\Phi(\alpha z) = \frac{2}{\sqrt{\pi}} \sum_{m=0}^{\infty} \frac{(-1)^m (2\alpha z)^m}{\Gamma(m+1)\Gamma(m+2)} \int_0^{\infty} x^{\frac{m}{2}} e^{-x}\, dx.$$

As is well known [9]

$$\Gamma\left(\frac{m}{2}+1\right) = \int_0^{\infty} x^{\frac{m}{2}} e^{-x}\, dx.$$

Consequently

$$\Phi(\alpha z) = \frac{2}{\sqrt{\pi}} \sum_{m=0}^{\infty} \frac{(-1)^m (2\alpha z)^m}{\Gamma(m+1)\Gamma(m+2)} \Gamma\left(\frac{m+2}{2}\right).$$

Using the fundamental property of the Γ-function [9]

$$\sqrt{\pi}\,\Gamma(m+2) = 2^{m+1}\,\Gamma\left(\frac{m+2}{2}\right)\Gamma\left(\frac{m+2}{2}+\frac{1}{2}\right),$$

we obtain

$$\Phi(\alpha z) = \sum_{m=0}^{\infty} \frac{(-1)^m (\alpha z)^m}{\Gamma(m+1)\,\Gamma\left(\frac{m+3}{2}\right)}.$$

REFERENCES

1. Ya. S. Itskhoki, Formirovaniye korotkikh impul'sov s pomoshch'yu magnitodielektrichesk linii, (Shaping of Short Pulses with the Aid of a Magnetodielectric Line) , 1947.
2. V. Carley, The Video Delay Lines of High ρ, Electronics, 1953, 4, 4, 26.
3. H.E. Callman, The Cable of High Impedance, Proceedings IRE, 1946, 34, 6.
4. S.I. Viglin, Investigation of Magnetodielectric Line, Tr. KhVAIVU, 1958, 106.
5. Ya. S. Itskhoki, Impulsnye ustroystva, Pulse Circuits, Izd. Sovetskoye Radio, 1959.
6. S. I. Viglin, Impulsnye ustroystve, Pulse Circuits, KhVAIVU, 1960.
7. L.I. Sirotinskiy, Overvoltages and Protection Against Overvoltages in Electric Apparatus, GEI, 1932.
8. V.A. Ditkin, P.I. Kuznetsov, Spravochnik po operatsionnomu ischesleniyu, Handbook of Operational Calculus, GITTL, 1951.
9. N.N. Lebedev Spetsial'nye funktsii i ikh prilozheniya, Special Functions and Their Applications, GITTL, 1953.

Submitted to the editor 27 October 1960

STATIC CHARACTERISTICS OF A PLANE MAGNETRON

A.B. Lazarev and G.F. Filimonov

The static characteristics of a plane magnetron are described in rigorous form. An expression is derived for the distribution of the electron density in the space-charge cloud. The values of the temperature of the electron gas, of the tangential current, and of the anode current of the magnetron are determined for $H > H_{cr}$. The theoretical results are in good agreement with the experimental data.

INTRODUCTION

Many theoretical papers have been published by now on the static characteristics of the magnetron. They differ in the approach to the problem, and can be divided into three groups.

In the first group of papers (see, for example [1, 2]), the main simplifying assumption is that the electron velocities on the cathode surface are equal. In these papers, analytic expressions for the distribution of the electron-space-charge density and for the electron velocities are derived, and the critical magnetic field of the magnetron H_{cr} is correctly calculated. The main shortcoming of references [1, 2] is the assumed single-velocity mode on the cathode surface, which clearly does not agree with reality and does not enable us to explain many essential factors in the static characteristics of the magnetron (for example, the existence of a residual anode current flowing through the magnetron when $H > H_{cr}$.

The second group of papers is connected with the solution obtained by Brillouin (see [3]) according to which there can be adjacent to the cathode of the magnetron an electron cloud of constant density, in which the electrons describe circular trajectories (concentric with the cathode), then an electron emitted from the cathode with arbitrary velocity and moving in static fields will return to the cathode after a definite time, or will settle on the anode. It is thus inexplicable how rotating electrons concentric with the cathode can be produced in a magnetron with hot cathode and form a Brillouin cloud. Furthermore, calculations made for several real magnetrons have shown that the radius of the boundary of the Brillouin cloud is much less than the radius of the cathode in these magnetrons, and no Brillouin conditions can be established at all in these magnetrons.

In the third group of papers (see [4, 5]) the solution of our problem is sought by methods of Maxwellian statistics. In [4], the distribution of the electrons is described by the Boltzmann therm dynamic equilibrium formula, using the free-energy potential (see Section 18.2 of reference [6]), and reference [5] even makes use of diffusion theory.

As is well known, the mean free path of the electrons in a magnetron is much greater than the magnetron dimensions. Therefore, electron collisions cannot play any noticeable role in either the formation of the electron cloud (static mode) or in the conversion of their energy into the energy of the high-frequency field (dynamic mode), and the attempt in [4, 5] to describe the operation of the magnetron is patently not justified. Experiment actually confirms that the main parameters of the magnetron (efficiency or output power) are independent of the degree of vacuum in the device, starting with vacuum on the order of $10^{-6} - 10^{-7}$ mm Hg.

The agreement between theory and experiment, obtained in [4, 5], is made possible only by free manipulation of arbitrary parameters, such as the volume temperature of the electrons in [4] or the volume temperature of the electrons and the diffusion coefficient in [5].

In the present paper we attempt a new approach to the solution of the problem of the static mode of the magnetron. We assume here that the cathode of the magnetron emits electrons with a certain velocity distribution function $F_c(v_0)$, which is identified with the distribution function of the thermionic electrons on the surface of the cathode, and which is assumed to be

$$F_c(v_0) = \frac{2n_0}{\pi v_T^2} e^{-\left(\frac{v_0}{v_T}\right)^2}, \tag{1}$$

where $v_T = \sqrt{2kT_c/m}$, T_c is the cathode temperature, k is Boltzmann's constant, m the electron mass, and n_0 the density of the emitted electrons on the cathode surface. The motion of the electrons in the interaction space is determined by the focusing fields, compared with which the space-charge field is assumed to be small.

These assumptions enable us to calculate the macroscopic characteristics of the electron

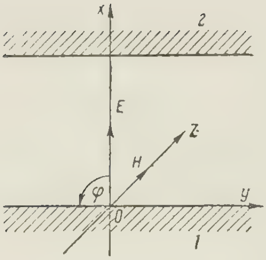


Figure 1. Schematic representation of plane magnetron: 1 — Cathode; 2 — Anode.

cloud in the magnetron: the particle volume density N = N(x), the particle current $\vec{j} = \vec{j}(x)$, and the temperature T = T(x) (see Figure 1).

The results obtained were checked on an experimental cylindrical magnetron with a wide cathode and smooth anode block (r_c = 1.6 mm, r_a = 2.4 mm), in which the focusing electric field deviated by about 25%.

1. FUNDAMENTAL RELATIONS

An investigation of the static mode of a magnetron should result in the evaluation of integrals of the type

$$I_{j,n} = \int v_j^n F(\vec{v})\, d\vec{v} = \int\limits_{C_T} v_j^n F_c(\vec{v}_0) \left| \frac{\partial(v_x, v_y)}{\partial(v_{0x}, v_{0y})} \right| d\vec{v}_0, \tag{2}$$

which determine N(x) (n = 0); $\vec{j}$(x) (n = 1); T° (n = 2). For this purpose it is necessary to deduce from the equation of motion

$$\frac{d^2\vec{r}}{dt^2} = \frac{e}{m}\left\{\vec{E}_0 + \frac{1}{c}\left[\frac{d\vec{r}}{dt}, \vec{H}_0\right]\right\} \tag{3}$$

the explicit relationship $\vec{v} = \vec{v}(\vec{v}_0, \vec{r}_0)$ and the region of integration C_T. In Equation (3), $\vec{E}_0$ and $\vec{H}_0$ are the focusing fields, $\vec{r}$ is the radius vector of the electron emitted from the cathode with initial velocity $\vec{v}_0$, $\vec{v} = d\vec{r}/dt$, e and m are the charge and mass of the electron, and c is the velocity of light. The motion of the electron is assumed plane.

The solution (3) is expediently presented in the form

$$\vec{r} = \vec{r}_0 - \gamma \vec{y}_0 t + \frac{\rho}{\omega_H},$$

$$\rho = (\vec{v}_0 + \gamma\vec{y}_0)\sin\omega_H t + [\vec{v}_0 + \gamma\vec{y}_0, \vec{z}_0](1 - \cos\omega_H t), \tag{4}$$

where $\omega_H = \frac{eH_0}{mc}$; $\gamma = \frac{E_0}{H_0}c$; $\vec{x}_0, \vec{y}_0, \vec{z}_0$, are unit vectors. It follows from (4) that the greatest distance that an electron with initial velocity v_0 moves away from the cathode is

$$\rho_{x\,\mathrm{extr}} = \omega_H x_{\mathrm{extr}} = \frac{\omega_H}{|\omega_H|}|v_0 + \gamma y_0| + v_{0y} + \gamma. \tag{5}$$

According to (1), $F_C(v_0)$ has a maximum when $v_0 = 0$. If we want electrons with $v_0 = 0$ to fall into the interaction space (x > 0) of the magnetron and to be able to interact with a high frequency, we must stipulate that

$$\frac{\gamma}{\omega_H} \simeq eE > 0. \tag{6}$$

We shall assume below that this condition is satisfied. From (4) we obtain also the followi expressions for v_x and v_y:

$$v_x = v_x(\vec{v}_0, \rho_x) = \pm\sqrt{v_{0x}^2 + 2\rho_x\left(v_{0y} + \gamma - \frac{\rho_x}{2}\right)},$$
$$v_y = v_y(\vec{v}_0, \rho_x) = v_{0y} - \rho_x, \tag{7}$$

where $+|v_x|$ corresponds to electrons moving away from the cathode ("taking off") while $-|v_x|$ corresponds to electrons falling on the cathode. Formulas (7) enable us to find the expression for the Jacobian of the transition

$$\left|\frac{\partial(v_x, v_y)}{\partial(v_{0x}, v_{0y})}\right| = \frac{v_{0x}}{|v_x|}. \tag{8}$$

The quantity ρ_x in (7) and (8) is the coordinate of the point of observation and is assumed fixed.

The integrals (2) are functions of x. Contributing to their values are those electrons, for which

$$\infty \geqslant x_{\text{extr}} \geqslant x; \quad x_a \geqslant x_{\text{extr}} \geqslant x \tag{9}$$

and which pass through the point x during takeoff or return to the cathode, respectively. Expanding the first inequality in (9), we can show that it is equivalent to

$$\begin{gathered}
0 \leqslant v_0 \leqslant |\gamma| - \frac{|\rho_x|}{2}, \quad -\frac{\pi}{2} \leqslant \varphi\varepsilon(\omega_H) \leqslant \frac{\pi}{2}, \\
0 \leqslant |\gamma| - \frac{|\rho_x|}{2} \leqslant v_0 \leqslant \infty, \frac{\pi}{2} \leqslant \varphi\varepsilon(\omega_H) \leqslant \arcsin\frac{\sqrt{v_0^2 + 2\gamma\rho_x} - |\rho_x|}{v_0}, \\
0 \leqslant \frac{|\rho_x|}{2} - |\gamma| \leqslant v_0 \leqslant \infty, \quad -\frac{\pi}{2} \leqslant \varphi\varepsilon(\omega_H) \leqslant \arcsin\frac{\sqrt{v_0^2 + 2\gamma\rho_x} - |\rho_x|}{v_0}. \\
\varepsilon(z) = \pm 1; \quad z \gtrless 0.
\end{gathered} \tag{10}$$

The second inequality in (9) yields the region of integration for electrons falling on the cathode. It is found to be equal to $G(x) - G(x_a)$, where $G(x)$ is the region of variation of v_0 defined in (10).

We can now rewrite the integrals in (2) as

$$I_{j,n} = \left[\int\limits_{G(x)} \frac{1 + \varepsilon(v_x)}{2} + \int\limits_{G(x) - G(x_a)} \frac{1 + \varepsilon(v_x)}{2}\right] v_j^n F_c(v_0) \frac{v_{0x}}{|v_x|} v_0 dv_0 d\varphi. \tag{11}$$

If v_j^n is independent of the sign of v_x (11) assumes a more compact form:

$$I_{j,n} = \int\limits_{2G(x) - G(x_a)} v_j^n F_c(v_0) \frac{v_{0x}}{|v_x|} v_0 \, dv_0 \, d\varphi. \tag{12}$$

2. ELECTRON DENSITY DISTRIBUTION

Putting n = 0 in (12), integrating with respect to φ and integrating the resultant expression by parts, we obtain the following expression for N(x):

$$\begin{gathered}
N(x) = I_{j,0}(x) = \frac{2n_0}{\pi} \int\limits_0^\infty du e^{-(u^2+\Gamma)^2} \frac{2|\gamma| + v_T(u^2 + \Gamma)}{2\gamma\rho_x + v_T^2(u^2 + \Gamma)^2} - \Delta_a N(x), \\
\Gamma = \frac{|\rho_x| - 2|\gamma|}{2v_T}.
\end{gathered} \tag{13}$$

The term $\Delta_a N(x)$ is due to the settling of the electrons on the anode and can be neglected in the transcritical mode ($H > H_{cr}$) at distances sufficiently remote from the anode, where

$$N(x) \gg N(x_a),$$

Let us consider the relative order of magnitude of the quantities in the integrand of (13). The factor $\exp[-(u^2 + \Gamma)^2]$ reaches a maximum when $u^2 = (|\Gamma| - \Gamma)/2$ and decreases rapidly when u^2 deviates from this value. The half-width of the curve $\exp[-(u^2 + \Gamma)^2]$ does not exceed unity. It is easy to see that within these limits the fraction under the integral sign in (13) changes into an infinitesimal of first order of smallness. Indeed $|\gamma|$ is close to the velocity of the slowed-down wave in the magnetron and its order is $|\gamma| \simeq (10^{-1} - 10^{-2})$, whereas v_T, corresponding to $T_c = 1{,}000$°K, has a value $v_T = 10^{-3}\,c$. We can therefore expect that formal expansion of the fraction under the integral sign in (13) in powers of $u^2 + |\Gamma| - \Gamma/2$ would enable us to represent the integral by a series in powers of the small parameter $v_T/2|\gamma|$.

After carrying out this transformation, we obtain

$$N(x)=\frac{2n_0}{\pi}\sqrt{2|\rho_x|v_T}\,\frac{2|\gamma|+v_T\frac{\Gamma+|\Gamma|}{2}}{2\gamma\rho_x+\left(v_T\frac{\Gamma+|\Gamma|}{2}\right)^2}\,\psi_{1/4}(\Gamma)\times$$

$$\times\left\{1-a(\Gamma)\frac{\psi_{1/4}(\Gamma)-\psi_{3/4}(\Gamma)}{2\psi_{1/4}(\Gamma)}-\frac{b}{4}\left[1+2\Gamma^2\frac{\psi_{1/4}(\Gamma)-\psi_{3/4}(\Gamma)}{\psi_{1/4}(\Gamma)}\right]-\ldots\right\},$$

$$a=\frac{v_T}{2|\gamma|+\frac{\Gamma+|\Gamma|}{2}}\left(1-2v_T\frac{\Gamma+|\Gamma|}{|\rho_x+2\gamma|}\right), \tag{14}$$

$$b=\frac{v_T^2}{2\gamma\rho_x+\left(v_T\frac{\Gamma+|\Gamma|}{2}\right)^2}\left[1+2v_T\frac{\Gamma+|\Gamma|}{|\rho_x+2\gamma|}\left(1-\frac{4v_T\Gamma}{|\rho_x+2\gamma|}\right)\right],$$

$$\psi_p(\Gamma)=\frac{1}{2}\sqrt{\frac{|\Gamma|}{2}}\,e^{-\frac{\Gamma^2}{2}}\left[K_p\left(\frac{\Gamma^2}{2}\right)+\pi\frac{1-\varepsilon(\Gamma)}{\sqrt{2}}I_p\left(\frac{\Gamma^2}{2}\right)\right], \tag{15}$$

$$\psi_{1/4}(\Gamma)=\int_0^\infty du\,e^{-(u^2+\Gamma)^2}.$$

It is easy to verify that when $\rho_x \gtrsim v_T$, the terms in the braces of Equation (14) form a decreasing sequence of numbers, with relative order $v_T/2\gamma$.

The course of the curve N = N(x) can be readily visualized by substituting in (14) the asymptotic expressions for the Bessel functions. Then, retaining the first term in (14) we obtain in the lowest approximation in powers of v_T

$$\frac{N(x)}{2n_0}=\begin{cases}1; & \rho_x=0; \qquad\qquad (16a)\\ \dfrac{v_T}{\sqrt{\pi|\rho_x||\rho_x-2\gamma|}}; & v_T\lesssim|\rho_x|\leqslant 2|\gamma|-v_T; \qquad (16b)\\ 0.243\sqrt{\dfrac{v_T}{|\gamma|}}; & \rho_x=2|\gamma|; \qquad (16c)\\ \sqrt{\dfrac{2}{\pi}\left|\dfrac{\rho_x}{\rho_x-2\gamma}\right|\dfrac{v_T}{|\rho_x+2\gamma|}}\,e^{-\Gamma^2}; & 2|\gamma|+v_T<|\rho_x|<|\rho_{xa}|. \qquad (16d)\end{cases}$$

It follows from (16) that as the distance from the cathode increases, the electron density decreas and reaches a minimum when $\rho_x = \gamma$. At distances $\rho_x = 2\gamma$ the density N(x) has a relative maximum, beyond which N(x) diminishes very rapidly on the anode side.

Thus, the boundary of the space-charge cloud lies near the value

$$\rho_x=2|\gamma|. \tag{17}$$

If $2|\gamma| < |\rho_{xa}|$, then the greater part of the electrons emitted from the cathode will strike the anode and large current will be produced in the anode circuit. On the other hand, if $2|\gamma| > |\rho_{xa}|$, then only the "tail" of the distribution N(x) will reach the anode, and the current in the anode will be small. Therefore, the equality

$$2|\gamma|=\rho_{xa} \tag{18}$$

defines the critical mode of the magnetron.

It is interesting to note that the formula (16b), which determines the variation of N(x) at the center of the space-charge cloud, is proportional to $|v_{0x}(0,\rho_x)|^{-1}$ [see (7)], and coincides with the distribution of the space charge on the surface of a cathode with zero initial velocity.

Figure 2 shows a plot of ρ = eN(x) for the operating mode of the investigated tube; n_0 was determined from the formula (see reference [7])

$$n_0 = N_x \sqrt{\frac{kT_c}{2\pi m}}\, e^{-\frac{e\varphi}{kT_c}}, \tag{19}$$

where N and φ are the density of the electrons in the conduction band of the oxide and the work function of the outside cathode, respectively; these were assumed to be $N_x = 4 \cdot 10^{18}$ and $\varphi = 1.3$ ev.

The experimental verification of the radial distribution of the electrons in the magnetron, undertaken in [8], has shown that in the center of the working volume of the magnetron there actually exists a space-charge density. Its relative value, which is somewhat greater than that shown in Figure 2 (see Figure 16c of reference [8]), is fully explained by the fact that the density measured in [8] concerned only electrons of energy greater than 20 ev.

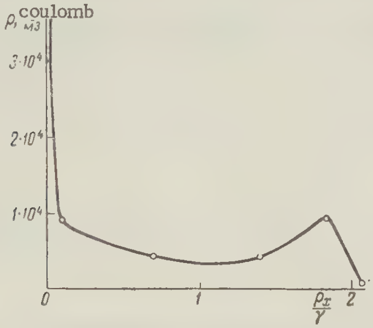


Figure 2. Distribution of density of space charge in magnetron ($\gamma = 3 \cdot 10^8$ cm/sec, $T_c = 1{,}000^\circ$ C, $r_c = 1.6$ mm, $r_a = 2.4$ mm).

Formulas (15) and (16) enable us to calculate the magnetron space-charge field E_q. Assuming that in the simplest case that the value of N(x) is given by (16b) when $|\rho_x| < 2|\gamma|$ and is zero when $|\rho_x| > 2|\gamma|$, we obtain the following rough estimate:

$$E_q \simeq \frac{8\pi^2 e n_0 v_T}{|\omega_H|\sqrt{\pi}}\left[\frac{\gamma}{\rho_{xa}} + \frac{1-\varepsilon(\Gamma)}{2}\left(\frac{2}{\pi}\arcsin\sqrt{\frac{\rho_x}{2\gamma}} - 1\right)\right]. \tag{20}$$

Substituting in (20) the numerical values of n_0, v_T, γ, and ρ_{xa} for the case shown in Figure 2, we find that the maximum of E_q is $9.4 \cdot 10^4$ v/m and amounts to only 10 per cent of the focusing field E_0.

3. TANGENTIAL VELOCITIES AND CURRENT

Multiplying (12) by e, putting $v_j^n = v_y$, and transforming the resultant expression in the same manner as the integral for N(x), we obtain the following expression for $j_y(x)$:

$$j_y(x) = -\varepsilon(\omega_H)\frac{4en_0\sqrt{2|\rho_x|}}{\pi v_T^2}\int_0^\infty du \frac{\left(\gamma + \frac{\rho_x}{2}\right)^2 - u^4}{[(u^4 + v_T\Gamma)^2 + 2\gamma\rho_x]^{3/2}} \times$$

$$\times \int_{|u^2+v_T\Gamma|}^\infty e^{-\frac{v^2}{v_T^2}} d\frac{(v^2+2\gamma\rho_x)^{3/2}}{3} - \Delta_a j_y \simeq \varepsilon(\omega_H)\frac{2en_0\sqrt{2|\rho_x} v_T}{\pi}\psi_{1/4}(\Gamma). \tag{21}$$

The second equality in (21) is valid when $v_T \lessapprox |\rho_x|$ and $N(x) \gg N(x_a)$, when the electrons settling on the anode have a negligible effect on the value of $j_y(x)$. Dividing the second equality in (21) by (14) in its lowest approximation (in powers of v_T), we obtain

$$v_y(x) = -\frac{2\gamma\rho_x + \left(v_T\frac{\Gamma + |\Gamma|}{2}\right)^2}{2|\gamma| + v_T\frac{\Gamma + |\Gamma|}{2}}\varepsilon(\omega_H). \tag{22}$$

We see from (21) and (22) that as T_c increases the average tangential current increases and the average tangential velocity remains constant at $\rho_x < 2|\gamma|$ (and equal to $v_y(0, \rho_x)$, (see 7).

The next approximation yields for $v_y(x)$ the correction

$$\frac{v_y - (v_y)_0}{(v_y)_0} = v_T\frac{2|\rho_x| + v_T(|\Gamma| - \Gamma)}{|\rho_x^2 + 2\gamma\rho_x|}|\Gamma|\frac{\psi_{1/4}(\Gamma) - \psi_{3/4}(\Gamma)}{2\psi_{1/4}(\Gamma)} =$$

$$= \begin{cases} \dfrac{v_T^2}{2\,|\rho_x^2 - 2\gamma\rho_x|}; & v_T \lesssim |\rho_x| < 2\gamma - v_T; \\ -\dfrac{v_T}{6\gamma}; & \rho_x = 2\gamma; \\ -\dfrac{v_T}{|\rho_x^2 - 4\gamma^2|} & 2\gamma + v_T < |\rho_x| < |\rho_{xa}|, \end{cases} \tag{23}$$

where $(v_y)_0$ denotes the right half of (22). Substituting in (23) the numerical values corresponding to T_c = 800 — 1000 degrees C, we can readily verify that the greatest change in velocity, due to a 10 per cent increase in T_c, does not exceed 1 per cent.

4. ANODE CURRENT

Of all the macroscopic quantities that characterize the magnetron space-charge cloud, the easiest to observe experimentally is the electron current in the anode. By calculating this current with the aid of (11) $(v_j^n = v_x > 0)$ we obtain

$$j_x^+(x) = \frac{en_0 v_T}{\pi}\left\{\sqrt{\pi}\,\frac{1-\varepsilon(\Gamma)}{2} + \varepsilon(\Gamma)\int_{|\Gamma|}^{\infty} e^{-v^2}dv + \int_{\sqrt{\Gamma^2 + \frac{2\gamma\rho_x}{v_T^2}}}^{\infty} e^{-v^2 + \frac{2\gamma\rho_x}{v_T^2}}dv\right\}. \tag{24}$$

When $\Gamma \gg 1$ this formula simplifies greatly and the residual current flowing in the magnetron anode in the transcritical mode is given by the formula

$$j_x^+(r_a) \simeq \frac{en_0 v_T}{\pi}\,\frac{2v_T\,|\rho_{xa}|}{|\rho_{xa}^2 - 4\gamma^2|}\,e^{-\Gamma^2}. \tag{25}$$

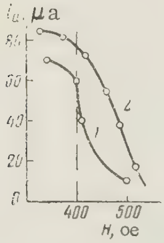


Figure 3

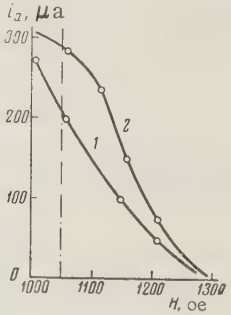


Figure 4

Figure 3. Dependence of anode current of magnetron, i_a, on the magnetic field intensity H with U_a = const.

1 — Calculated by formula (24) ; 2 — plotted from the experimental data (U_a = 100 v, T_c = 100 degrees C, r_c = 1.6 mm, r_a = 2.4 mm) .

Figure 4. Dependence of anode current on magnetron, i_a, magnetic field intensity H, U_a = = const.

1 — Calculated by formula (24) ; 2 — constructed from experimental data (U_a = 400 v T_c = 1,000 degrees C, r_c = 1.6 mm, r_a = 2.4 mm) .

The dependence of the anode current of the magnetron on the magnetic field, in a mode close to critical, is shown by the curves of Figures 3 and 4. Curves 1 were calculated from formula (24), while Curves 2 are based on the experimental data obtained with the aforementioned model with smooth anode. The experimental error did not exceed ±10 per cent, and the errors were essentially due to inaccuracies in the measurement of the magnetic field.

The agreement between the theoretical and experimental data is good; a certain discrepancy is apparently due to the fact that the approximate value of N_χ was used in the calculations.

At low currents (on the order of several microamperes) the slope of the theoretical curve is twice the slope of the experimental characteristic. Additional current sources are the volume ionization of the residual gases and secondary emission of electrons from the cathode

5. ELECTRON TEMPERATURE

Experiment has shown (see [8]) that the temperature of the neutral gas in the center of the working space of the magnetron is quite high and much greater than T_c. This effect can be brought about only by interaction between the neutral gas and the electron current, and gives certain grounds for assuming that the electron temperature at the center of the magnetron is also greater than T_c. The same conclusion was arrived at by the authors of references [4, 5] in their attempts to reconcile the curves derived in [4, 5] for the anode current with experiment. In view of the fact that the electron collisions cannot raise their temperature appreciably, it is natural to assume that the increase in the mean-square spread of the electron velocities (which is equivalent to their "heating") is the result of the motion of the electrons in external electromagnetic fields, and is of purely kinematic origin.

Indeed, calculating the mean kinetic energy of the electrons

$$\overline{\frac{mv^2}{2}} = m\frac{I_{x,2}(x)+I_{y,2}(y)}{2N(x)} \simeq kT_c + \frac{m}{2}\left[2\gamma\rho_x + \left(v_T\frac{\Gamma+|\Gamma|}{2}\right)^2\right] \tag{26}$$

and using (22) [the second equation in (26) is valid when $v_T \lessapprox |\rho|$, $N(x) \gg N(x_a)$], we obtain the following expression for the temperature of the electron gas

$$T = \frac{2}{3k}\left[\frac{m}{2}(\overline{v^2} - \overline{v}^2) + \frac{kT_c}{2}\right] = T_c + \frac{mv_T|\rho_x|}{3k}\,\frac{|\Gamma|-\Gamma}{2\gamma}\,. \tag{27}$$

Formula (27) shows that the electron temperature varies quadratically in the direction of the anode, and reaches a maximum $T_{max} = T_c(1 + 2|\gamma|/3v_T)$ when $\rho_x = \gamma$. Beyond the boundary of the space-charge cloud, the temperature of the electrons coincides with T_c.

In a stationary system of coordinates [see (26)] the average kinetic energy of the electron increases monotonically in the direction of the anode. We note that it is precisely this energy that determines the interaction between the electrons and the neutral gas.

CONCLUSION

In the present paper we described rigorously (without added assumptions and without introducing arbitrary coefficients) the static mode of a plane magnetron. An expression was derived for the distribution of the electron density in the space-charge cloud, from which follows that the electron density has a maximum near the boundary of the cloud. The theoretical dependence of the anode current of the magnetron in the critical mode on the magnetic field is in good agreement with the experimental data and explains the presence of a residual current. Comparison of the theoretical and experimental data enables us to estimate the influence of secondary current sources (volume ionization of the gas and secondary emission from the cathode) on the magnitude of the anode current.

The calculated values of the temperature of the electron gas coincide in order of magnitude with its qualitative estimate, obtained from the noise characteristics of an under-excited magnetron.

As shown by theoretical and experimental researches (see, for example, [9,10]), the space-charge cloud is unstable against high-frequency oscillations. However, these oscillations are of low intensity and produce a high-frequency field which does not exceed 0.1 per cent of the static field E_0. Therefore, the deviation that these oscillations introduce into the mean value of the space-charge density, given by formula (14), likewise does not exceed 0.1 per cent.

REFERENCES

1. G.A. Grinberg, Izbrannye voprosy matematicheskoy teorii elektricheskikh i magnitnykh yavleniy (Selected Problems in the Mathematical Theory of Electric and Magnetic Phenomena), Chapter VI, Izd. AN SSSR, 1948.

2. S. Ya. Braude, ZhETF, 1935, 5, 621; 1940, 10, 217.
3. Teoriya magnetrona (po Brillyuenu) [Magnetron Theory (after Brillouin)] Collection of Translations from the English, Izd. Sovyetskoye radio, 1946.
4. G. Ya. Lyubarskiy, L.E. Pargamannik, Dokl. AN SSSR, 1952, 34, 491.
5. L.E. Pargamannik, M. Ya. Mints, ZhTF, 1957, 27, 1301.
6. S. Chapman and T.G. Cowling, Mathematical Theory of Nonuniform Gases, Cambridge, 1953.
7. G. Hermann and S. Wagener, The Oxide Coated Cathode, Chapman and Hall, London. 1951.
8. H.C. Neddermann, Journal of Applied Physics, 1955, 26, 1420.
9. B.B. Kadomtsev, ZhTF, 1959, 29, 831.
10. V.P. Tychioskiy, Radiotekhnika i elektronika, 1956, 1, 2, 233; 1956, 1, 3, 344.

Submitted to the editors, 4 July, 1960.

POSSIBILITY OF EXISTENCE OF SLOW WAVES IN AN ISOTROPIC PLASMA WAVEGUIDE

G.A. Postnov

The problem of propagation of a magnetic wave in a plane plasma waveguide with dielectric insert in the middle is solved. It is shown that the inhomogeneity of the plasma filling the waveguide makes possible the existence of slow waves in the guide.

Slow guided waves, i.e., waves whose wave number in the waveguide is greater than the wave number in the free space are used in many fields of microwave electronics. These can be produced by filling the waveguide with a dielectric, by placing a ridged structure in the waveguide, etc. One of the methods of slowing down the waves is to fill the waveguide with an anisotropic material. As shown in [1], a wave can be slowed down by placing transverse to the axis of a cylindrical waveguide a stratified dielectric in which layers of thickness a with dielectric constant ϵ alternate with layers of thickness b and dielectric constant unity (in general, ϵ_1). The transverse wave number of such a waveguide is given by the formula

$$g^2 = \frac{\varepsilon_r}{\varepsilon_z}(k^2\varepsilon_z - h^2), \tag{1}$$

where

$$\varepsilon_r = \frac{a\varepsilon + b}{a + b}; \quad \varepsilon_z = \frac{(a + b)\varepsilon}{a + b\varepsilon}.$$

It is easy to see that the slow waves, for which $g^2 < 0$, can be obtained even if the dielectric constant is less than unity. This will take place, for example when $\epsilon < -b/a$ and $-1 < \epsilon < 0$.

Another widely known method of slowing down waves in a guide is to fill the guide with a gyrotropic plasma. This method calls for the production of appreciable magnetic fields particularly in the centimeter band, so that its practical use is quite difficult. However, as will be shown presently, slow waves can apparently be obtained by using also the inhomogeneity of the plasma inside the waveguide in the absence of a magnetic field. This inhomogeneity is brought about by the decrease in the electron concentration toward the walls of the waveguide as a result of diffusion. As can be seen from [2] although the phase velocity in a plane waveguide with plasma does decrease compared with a uniformly filled waveguide, no slow waves are formed. To obtain slow waves it is necessary to introduce into the waveguide still another surface, which changes the distribution of the electron concentration and consequently changes the permittivity of the plasma in the waveguide.

Let us examine the plane waveguide with plasma (Figure 1). If we place inside the waveguide, at equal distance from the walls, a thin dielectric plate with permittivity close to unity, it will not greatly distort the field in the waveguide, and the decrease in plasma concentration in the vicinity of the plate will be similar to that produced near a metallic wall. Approximating the distribution N between two surfaces by a cosine function, we can write for this case

$$N = N_0 \left| \sin \frac{\pi y}{2b} \right|, \tag{2}$$

where 2b is the distance between walls of the plane waveguide.

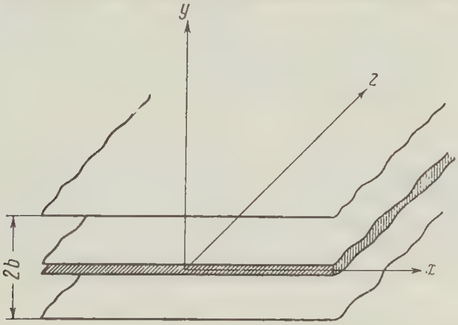


Figure 1

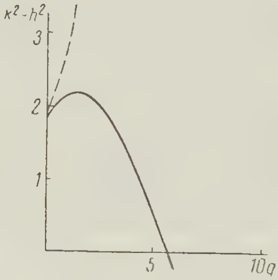


Figure 2

Figure 1. Plane waveguide with dielectric insert.
Figure 2. Dependence of transverse wave number on the concentration of the plasma in plane waveguide with magnetic wave. Solid line — waveguide with dielectric plate, dotted line — waveguide without insert.

Under this assumption, the wave equation for the magnetic waves assumes the form

$$\frac{d^2 E_x}{dy^2} + \left[k^2 - h^2 - k^2 \frac{4\pi e^2 N_0}{m\omega^2} \left| \sin \frac{\pi y}{2b} \right| \right] E_x = 0 \tag{3}$$

with boundary conditions $E_x \big|_{y=-b,\,b} = 0$. This is a Hill equation with period π. Let us expand Equation 2 in a Fourier series over the interval 2b:

$$\left| \sin \frac{\pi y}{2b} \right| = \frac{2}{\pi} - \frac{4}{3\pi} \cos 2 \frac{\pi y}{2b} - \frac{4}{15\pi} \cos 4 \frac{\pi y}{2b} - \ldots . \tag{4}$$

We see that the series of Fourier coefficients converges absolutely. Substituting (4) in (3) and putting $4\pi e N_0/mc^2 = 2q$, we rewrite (3) in the form

$$\frac{d^2 E_x}{d\zeta^2} + \left[\theta_0 + 2 \sum_{n=1}^{\infty} \theta_{2n} \cos 2n\zeta \right] E_x = 0. \tag{5}$$

Here

$$\zeta = \frac{\pi y}{2b}; \quad \theta_0 = \left(\frac{2b}{\pi} \right)^2 \left[k^2 - h^2 - \frac{4q}{\pi} \right]; \quad \theta_{2n} = \frac{4q}{\pi (4n^2 - 1)}.$$

The general solution of the Hill equation on the basis of the Floquet theorem can be written in the form

$$E_x = Ae^{\mu\zeta}\Phi(\zeta, \sigma) + Be^{-\mu\zeta}\Phi(\zeta, -\sigma),$$

where σ is a certain parameter [3]. The function Φ can also be represented in the form of a trigonometric series, where σ plays the role of the phase. To satisfy the boundary condition it is necessary that E_x be an even function. This means that $B = 0$. This condition is satisfied when $\sigma = -\pi/2$. In this case, $\mu = 0$ and θ_0 is an eigenvalue of Equation (5).

Using the formulas given in [3], we write

$$\Theta_0 = 1 - \Theta_2 - \frac{1}{8}\Theta_2^2 - \frac{1}{4}\Theta_2\Theta_4 - \frac{1}{6}\Theta_4^2 - \ldots \tag{6}$$

or, substituting the values of the coefficients,

$$\left(\frac{2b}{\pi}\right)^2\left[k^2 - h^2 - \frac{4q}{\pi}\right] = 1 - 0{,}4264q - 0{,}0783q^2 + 0{,}0001q^3 - \ldots .$$

A plot of $(k^2 - h^2)$ vs. q at a 3-cm wavelength is shown in Figure 2. For comparison, the same figure shows the same relation for a plane plasma waveguide without dielectric, calculated in accordance with reference [2]. We see that, starting with a value $q = 5.56$ slower waves begin to propagate in the waveguide. This corresponds to a concentration $N_0 = 4.24 \cdot 10^{12}$ cm^{-3}.

This conclusion is of certain interest, all the more since it follows from [1] that when the waveguide is filled with a substance with abrupt changes in ϵ, only electric slow waves can be obtained, but not magnetic ones. In addition, the method of obtaining slow waves in a plasma waveguide, by introducing only a single supplementary surface (in a round waveguide, probably, only a thin coaxial dielectric rod), results in relatively low losses and in some cases may prove to be more convenient than other methods of producing slow-wave structures.

I am deeply grateful to Ya. B. Faynberg for valuable hints.

REFERENCES

1. Ya. B. Faynberg, N.A. Khizhnyak, ZhTF, 1955, 25, 4, 711.
2. G.A. Postnov, Propagation of Waves in an Isotropic Plasma Waveguide. Radiotekhnika i elektronika, 1960, 5, 10, 1598.
3. N.W. McLachlan, Theory and Applications of Mathieu Functions, Clarendon Press, Oxford, 1947.

Submitted to the editor, 22 February, 1960.

CONCERNING THE SENSITIVITY OF RADIOMETERS

G.P. Apushkinskiy

The effect of fluctuations of gain on the sensitivity of a superboardband null radiometer is examined.

From the point of view of the effect of fluctuations of gain on the sensitivity of a radiometer, the compensation, modulation, and null methods are three successive stages of improvement.

In the compensation method [1, 2], the fluctuations of the gain are involved with a small "weighting factor," determined by the fact that the high level of internal receiver noise passes through the entire amplification channel.

In the modulation method [1, 3] only the noise signal passes through the entire amplification channel, and this signal is usually less than the apparatus noise (by a factor of n), and consequently gain fluctuations will have a lesser effect on the sensitivity (by an approximate factor of n). In particular, when the bandwidth is increased the limiting sensitivity, connected with the gain fluctuations, will be $n = (T_i + T_a)/\Delta T$ times higher for a modulation radiometer than for a compensation radiometer (see reference [3], p. 740). In this case, although the full level of internal noise prevails in the channel up to the detector preceding the narrow-band low-frequency amplifier, one can nonetheless neglect the associated effect of gain fluctuations under prevailing relationships between the gains and bandwidths before and after the detector.

In the null method [4, 3] there is no signal in the channel past the detector (i.e., in the low-frequency amplifier and beyond), and since the gain in this part of the channel constitutes the main gain of the radiometer (usually an increase in power of several million), the contribution of the gain fluctuations to the sensitivity is reduced to a negligible amount.

However, the use of traveling-wave tubes in radiometers [5], and in particular the development of the direct-amplification TWT radiometer [6] have made it possible to increase the bandwidth to 10 per cent. Suggestions have been made to increase the bandwidth of the radiometer to 20 per cent [7], inasmuch as the noise problem entailed thereby can be solved in principle. In this case, an analysis of the role of fluctuations in the sensitivity of a null radiometer may be of interest, The noise spectrum existing here past the detector (ahead of the low-frequency amplifier), due to the input temperature T (the intrinsic noise plus the load noise) and due to the gain fluctuations, has the following form (reference [2], p. 433):

$$G_\alpha(\Omega) = 4\overline{\alpha^2}T^2 \frac{\Delta\Omega_\alpha}{\Delta\Omega_\alpha^2 + \frac{\pi^2}{4}\Omega^2}, \tag{1}$$

where $\Delta\Omega_\alpha$ is the effective band of the spectrum of fluctuations of $\alpha(t)$.

The filter of the low-frequency amplifier picks out of this spectrum the harmonic of the modulation frequency. Past the filter, the mean-square value of the fluctuations is given by the formula

$$\overline{Y^2} = \int_{\Omega_1}^{\Omega_2} G_\alpha(\Omega)\,\Phi(\Omega)\,d\Omega, \tag{2}$$

where $\Delta\Omega_\alpha$ is the characteristic of the low-frequency-amplifier filter with passband in the interval $\Omega_1 - \Omega_2$.

In particular, for a rectangular filter band, we obtain

$$\overline{Y^2} = \frac{8\overline{\alpha^2}T^2}{\pi}\left[\operatorname{arc\,tan}\left(\frac{\pi\Omega}{2\Delta\Omega_\alpha}\right)\right]_{\Omega_1}^{\Omega_2}. \tag{3}$$

Usually the modulation frequency Ω_m (which lies in the vicinity of Ω_1 and Ω_2 and between the two) is chosen greater than $\Delta\Omega_\alpha$. Taking this into account, we can obtain approximately

$$\overline{Y^2} = \frac{16\,\overline{\alpha^2}T^2\Delta\Omega_\alpha\Delta\Omega_f}{\pi^2\Omega_m^2}, \tag{4}$$

where $\Delta\Omega_f$ is the effective bandwidth of the filter.

The synchronous detector, mixing the fluctuations with the reference voltage of frequency Ω_m, "shifts" the spectrum to the null region, and the integrating device that follows it usually has a much narrower bandwidth $\Delta\Omega$ than the low-frequency amplifier filter bandwidth $\Delta\Omega_f$ so that filtering is determined by $\Delta\Omega$, and we finally have

$$\overline{Y^2} = \frac{16\,\overline{\alpha^2}T^2\Delta\Omega_\alpha\Delta\Omega}{\pi^2\Omega_m^2}. \tag{5}$$

Further, determining the sensitivity of the radiometer in the usual fashion [2], we obtain in lieu of the known expression for the sensitivity of a null radiometer [3]

$$\delta T = \pi T\left(\frac{\Delta\Omega}{\Delta\omega}\right)^{1/2} \tag{6}$$

(where $\Delta\omega$ is the bandwidth of the radiometer, and the effect of the gain fluctuations is not taken into account at all), a new expression

$$\delta T = \pi T\left[\frac{\Delta\Omega}{\Delta\omega} + \frac{16\,\overline{\alpha^2}\Delta\Omega_\alpha\Delta\Omega}{\pi^2\Omega_M^2}\right]^{1/2}. \tag{7}$$

When the radiometer bandwidth $\Delta\omega$ is broadened and the power supply for the traveling-wave tube is insufficiently stabilized, both terms in the brackets may become of equal order of magnitude.

It is necessary to choose the modulation frequency Ω_m as high as possible without tending to increase the stabilizer supply frequency for the traveling-wave tube. Although the increase in frequency decreases the weight and dimensions of the apparatus, the effective band of the fluctuation spectrum $\Delta\Omega_\alpha$ is also increased thereby.

The recommendations given by formula (7) may be of interest also for a radiometer with a traveling-wave maser, when the band $\Delta\omega$ is not as wide as in a TWT radiometer, but then the gain fluctuations (i.e., $\overline{\alpha^2}$) may be considerably greater.

REFERENCES

1. V.S. Troitskiy, ZhTF, 21, 8, 994.
2. F.V. Bunkin, N.V. Karlov, ZhTF, 1955, 25, 3, 430.
3. F.V. Bunkin, N.V. Karlov, ZhTF, 1955, 25, 3, 733.
4. V.S. Troitskiy, ZhTF, 1955, 25, 3, 478
5. G.P. Apushkinskiy, Radiotekhnika i elektronika, 1958, 3, 6, 852.
6. F.D. Drake, H.I. Ewen, Proceedings IRE, 1958, 46. 1, 53,
7. N.A. Bol'shakov, Yu. N. Pariiskii Izv. GAO, 1960, 21, 5, No. 164, 162.

Submitted to the editor 22 December 1960

INVESTIGATION OF FIELD EFFECT TRANSISTORS

Z.S. Gribnikov, B.N. Kompan and L.V. Svyatogor

We describe the production of diffusion field-effect transistors made of germanium. The static parameters and the characteristics of such transistors are investigated under various operating conditions. Phototriode operation is considered and the photosensitivity of the transistors is determined and found to be 100 amp/lumen.

INTRODUCTION

Channel or field-effect transistors were proposed by Shockley [1] and first investigated by Dacey and Ross [2, 3]. They have a high maximum gain frequency and large input resistance, and can also operate as very sensitive phototriodes. The search for effective methods of production of such transistor and an investigation of their properties is therefore a timely problem.

1. ELEMENTS OF THE THEORY

1. Homogeneous Channel in the Approximation of a Smoothly Varying Potential.

We shall call a channel homogeneous if: a) the electric field in the channel depends on a single coordinate; b) the channel is characterized by a single value of the cutoff voltage (V_0), at which the layer depleted of carriers extends over the entire cross section of the channel. In the approximation where the potential varies smoothly along the channel, an approximation valid within limits established in reference [1], the volt-ampere characteristic of such a channel is given by the formula

$$i = [f(\varphi_d) - f(\varphi_{so})], \tag{1}$$

where $\varphi = V/V_0$; V is the voltage between some point of the channel and the gate; φ_d and φ_{so} are the values of φ pertaining to the drain and to the source of the channel, respectively*; $i = I/I_0$; $I_0 = G_0 V_0$, where G_0 is the conductance of the channel when $\varphi_{so} = 0$ and $\varphi_d \to 0$; $f(\varphi)$ is the so-called "channel function," the form of which depends on the geometry and on other properties of the gate and of the channel.

For a plane-parallel one-dimensional channel with sharp p-n junction and highly doped gate (for example, with p^+-n junction, formed by linear variation of the impurity concentration see [1]), we have

$$f(\varphi) = \varphi\left(1 - \frac{2}{3}\varphi^{1/2}\right); \tag{2}$$

For a plane-parallel one-dimensional channel with p-n junction formed by linear variation of the impurity concentration (see [4]), we have

$$f(\varphi) = \varphi\left(1 - \frac{3}{5}\varphi^{2/3}\right). \tag{3}$$

Formulas (2) and (3) are valid when $0 \leqslant \varphi \leqslant 1$.

When $\varphi_d = 1$, the channel is cut off. The field in the channel then becomes infinitely large near the drain. A further increase in φ_d does not change the distribution of the potential in the channel, except for a small section near the drain. This enables us to assume that $f(\varphi) = f(1)$ when $\varphi > 1$.

Recognising that when $\varphi > \varphi_e = V_e/V_0$ a breakdown of the p-n junction takes place, and that injection is observed when $\varphi < 0$, we can represent the channel function $f(\varphi)$ in the form plotted in Figure 1, and make use of formula (1) for arbitrary φ_d and φ_{so}. Using the plot of $f(\varphi)$ as a template, we can construct a family of characteristics $i = F(\varphi_d - \varphi_{so})$ for constant φ_{so} using arbitrary φ_d and φ_{so}. (This is done in Figure 1).

---

*We use here the same terminology as in the Russian translation of references [1, 2]; for channels with electron conductivity the drain is the anode, the source is the cathode and the gate is the grid.

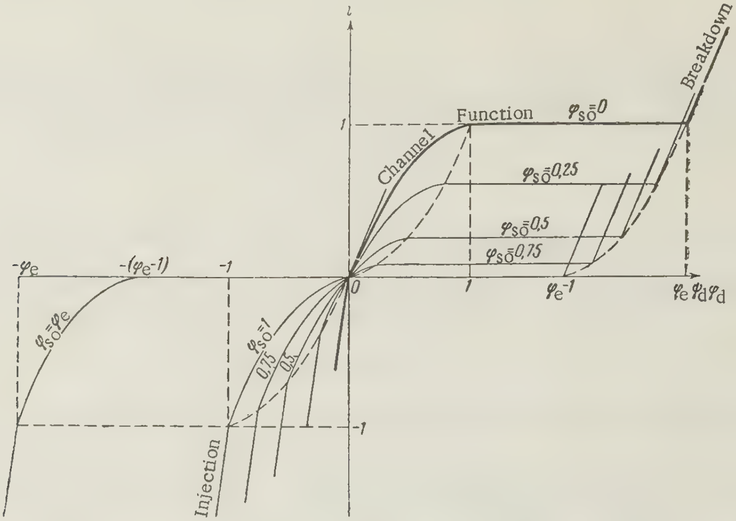


Figure 1. Family of characteristics $i = F(\psi_d)$ for different values of φ_{so}, constructed with the aid of the channel function $f(\varphi)$ (heavy curve); $\psi_d = \varphi_d - \varphi_{so}$.

The main static parameter of the field-effect transistor is the mutual conductance, defined as

$$G = \left(\frac{\partial I}{\partial V_{so}}\right)_{(V_d - V_{so}) = \text{const}} \tag{4}$$

In the cutoff mode, its value is

$$G = G_0 \left[\frac{df(\varphi_{so})}{d\varphi_{so}}\right]. \tag{4'}$$

When $\varphi_{dr} = 0$, we have $df(\varphi_d)/d\varphi_d = 1$ and $G = G_0$.

2. Account of the Finite Limiting Drift Velocity of the Carriers in the Electric Field

The cutoff criterion $E = \infty$ is true if the carrier drift velocity increases without limit as the electric field is increased. In fact [6] at fields greater than 10^3 v/cm this increase is greatly slowed down in germanium and in silicon and quasi-saturation sets in, the carriers reaching a limiting drift velocity (for electrons in germanium its value is $v_{lim} = 5.4 \cdot 10^6$ cm/sec). Assuming for simplicity that the carrier drift velocity is directly proportional to the intensity of the field, $v_d = u_0 E$, where u_0 is the mobility of the carriers in weak fields, so long as $v_d \leqslant v_{lim}$, and is then independent of the field, we introduce a saturation criterion in the form

$$E = E_c = \frac{v_{lim}}{u_0}. \tag{5}$$

Figure 2 shows the dependence of $v_d = f(E)$ for electrons in germanium, taken from reference [6], compared with the simplified relationship, used in the calculation. In this case $E_c = 1400$ v/cm.

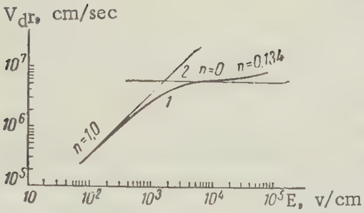


Figure 2. Dependence of the electron drift velocity in germanium on the electric field intensity. 1 — Experimental dependence at 300 degrees K [6] ; 2 — dependence used for the calculations.

The saturation voltage and current I_s and V_s of the field-effect transistor is determined, using the criterion (5) by solving the system

$$\varphi_c \frac{df}{d\varphi}\bigg|_{\varphi=\varphi_s} = i_s \qquad (6a) \qquad\qquad i_s = f(\varphi_s) - f(\varphi_s), \qquad (6b)$$

where $i_s = I_s/I_0$, $\varphi_s = V_s/V_0$, and $\varphi_c = E_c l/V_0$ [Equation (6b) is Equation (1) rewritten in the new notation, while Equation (6a) follows directly from condition (5)]. From the system (6a) and (6b) we obtain for φ_s

$$\varphi_c(1 - \varphi_s^m) = \varphi_s\left(1 - \frac{\varphi_s^m}{m+1}\right) - \varphi_s\left(1 - \frac{\varphi_{so}^m}{m+1}\right), \qquad (7)$$

where m = 1/2 when $f(\varphi)$ is given by formula (2), and m = 2/3 when $f(\varphi)$ is given by (3).

Figure 3 shows the dependence of the channel saturation voltage $\varphi_s - \varphi_{so}$ on φ_c for various values of φ_{so} (for m = 1/2). Unlike the limiting case $\varphi_c = \infty$, when $\varphi_s = 1$ no matter what the value of φ_{so}, the value of φ_s depends on φ_{so} when φ_c is finite. Therefore, when $\varphi_c < 10$, the family of characteristics can no longer be plotted by means of the template (Figure 1). The saturation current I_s becomes less than I_0 when $\varphi_{so} = 0$, and the average mutual conductance becomes less than $mG_0/(m + 1)$. Figure 4 shows the dependence of i_s (for $\varphi_{so} = 0$) on φ_c (m = 1/2).

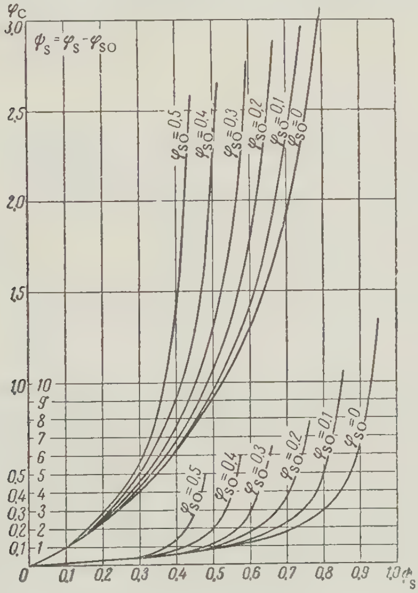


Figure 3. Dependence of $\psi_s = \varphi_s - \varphi_{so}$ on φ_c for different values of φ_{so} for channels with steep p-n gate junction.

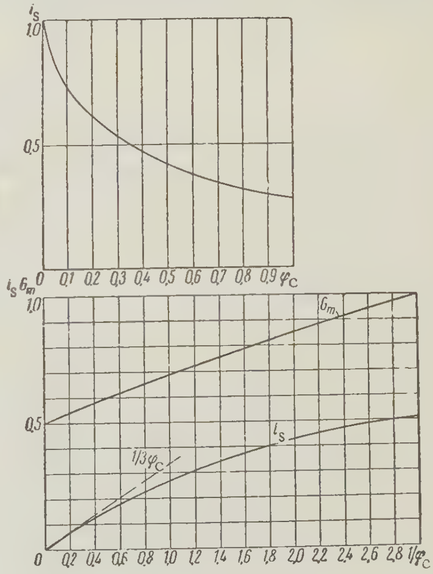


Figure 4. Dependence of i_s and G_m on φ_c with $\varphi_{so} = 0$ for channels with steep p-n junction barrier.

The mutual conductance of the transistor, defined by formula (4), can be obtained in the form

$$G=\left(m\varphi_c\varphi_s^{m-1}\frac{1-\varphi_{so}^m}{1-\varphi_s^m+m\varphi_c\varphi_s^{m-1}}\right)G_0. \tag{8}$$

When $\varphi_c \to \infty$, formula (8) gives the same result as (4'), namely $G = G_0(1 - \varphi_{so}^m)$. When $\varphi_c \to 0$ we get $\varphi_s \to \varphi_c$ and $G = m\varphi_c^m G_0(1 - \varphi_{so}^m)$.

Let us formulate our conclusions:

1) In long channels ($\varphi_c \geqslant 1.5 - 2$), the maximum current and the average mutual conductance are inversely proportional to the length of the channel $[I_s = mV_0G_0/(m + 1)]$. A further decrease in the length of the channel leads to a slowing down in the increase of I_s, which assumes at $\varphi_c \to 0$ a saturation value

$$I_s = G_0E_c l = g_0E_c, \tag{9}$$

where g_0 is the conductivity of a channel 1 cm long.

2) The saturation voltage for $\varphi_{so} = 0$ in long channels is equal to V_0. With decreasing channel length $V_s < V_0$ (for example, when $m = 1/2$, and $\varphi_c = 1$ we have $V_s = V_0/2$). When $\varphi_c \to 0$ we get $V_s = E_c l$.

3) The power released in the channel at saturation voltage diminishes slowly with increasing channel length.

4) When $\varphi_{so} \neq 0$, the value of V_s depends on V_{so}, increasing to V_0 as $V_{s0} \to V_0$.

3. Effect of Small Inhomogeneity

It is difficult to produce perfectly homogeneous channels. A small inhomogeneity can be accounted for in the theory of smooth variation of potential in the channel.

Assuming that the cutoff voltage and the conductivity of the channel vary along the channel so that

$$\frac{df}{d\varphi} = 1 - \varphi^m + \gamma x \tag{10}$$

[for the case of a steep p-n junction barrier, this corresponds to a width varying channel as $a(x) = a_0(1 + \gamma x)$]. The normalizing voltage V_0 in formula (10) is referred to the point $x = 0$. The volt-ampere characteristic is then obtained by integrating the equation

$$I = g_0V_0(1 - \varphi^m + \gamma x)\frac{d\varphi}{dx} \tag{11}$$

(g_0, like V_0 is referred to the point $x = 0$) and has the form

$$i = \int_{\varphi_s}^{\varphi_{dr}} (1 - \varphi^m)\exp\left[\frac{\gamma l}{i}(\varphi - \varphi_{so})\right]d\varphi, \tag{12}$$

where, as before, $i = I/I_0$, with $I_0 = g_0V_0/l$.

In the case of small inhomogeneity, we obtain from (12)

$$i = f(\varphi_{dr}) - f(\varphi_s) + \gamma l\left[\frac{f_1(\varphi_{dr}) - f_1(\varphi_{so})}{f(\varphi_{dr}) - f(\varphi_{so})} - \varphi_{\text{н}}\right], \tag{13}$$

where $f_1(\varphi) = \varphi^2/2\ (1 - 2/m + 2\cdot\varphi^m)$. Then asymmetry sets in and the plotting of the family of characteristics by means of the template is again impossible.

2. PREPARATION OF EXPERIMENTAL SPECIMENS

We used as experimental specimens end-connected diffusion field-effect transistors. The manufacture of unipolar transistors on the base of a diffused p-n junction is described in references [4, 5], but the methods proposed there call for very high mechanical precision if channels with sufficiently large perimeter are to be obtained. Sufficient accuracy is apparently unattainable, and the channel current is much less than the calculated value (see [5]). Therefore, to obtain specimens with high mutual conductance, we employed a procedure procedure in which the diffusion of the impurity and the etching occur on the same surface, and, consequently, there is no need for high accuracy in the mechanical working. The experimental specimens are shown schematically in Figure 5.

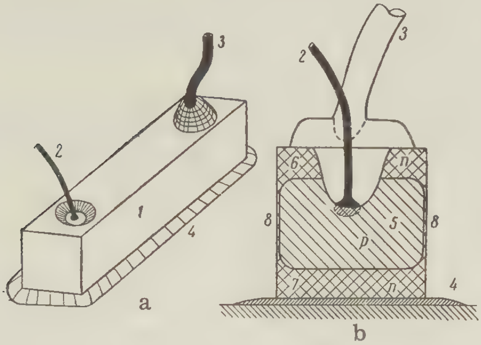


Figure 5. Experimental transistor specimens (a — general view, b — cross section). 1 — Germanium slab; 2 — lead from p-region of gate; 3 — source lead; 4 — holder, drain lead; 5 — gate region, 6 — source region; 7 — drain region, 8 — channel region.

3. INVESTIGATION OF EXPERIMENTAL SPECIMENS

The experimental specimens had channels of length $l \simeq 0.025$ cm, the extent of the front of the channel (perimeter) was $d \simeq 1$ cm. The conductivity of the channel was in the range $G_0 \simeq 0.002 - 0.005$ mho; the cutoff voltage V_0 was 15 — 40 v. From these data we can calculate the saturation current I_0, the thickness of the channel a, and the gradient of the impurity concentration in the cross section of the channel $\mathfrak{G}$, assuming linear variation of the concentration [4]:

$$I_0 = \frac{2}{5} G_0 V_0, \tag{14}$$

$$a = \frac{3\varepsilon u_n d}{16\pi l} \frac{V_0}{G_0}, \tag{15}$$

$$\mathfrak{G} = \frac{2lG_0}{qu_n da^2}. \tag{16}$$

We get $a \simeq 6 - 10$ microns and $\mathfrak{G} \simeq 10^{19}\ \mathrm{cm}^{-4}$. The approximate value of $\mathfrak{G}$ can be obtained from calculation of the diffusion process, by measuring the depth of the p-n junction:

$$\mathfrak{G} = \frac{F(C_0, \mathfrak{N}_p)}{x_{pn}}, \tag{17}$$

where $F(C_0, \mathfrak{N}_p)$ is a function of the concentration $\mathfrak{N}_p$ of the main impurity in the parent material, and of the concentration of the diffusing impurity at the surface C_0. The value of $\mathfrak{G}$ obtained from (17) is also about $10^{19}\mathrm{cm}^{-4}$.

1. Electric Characteristics and Parameters

The measurement of the statistical characteristics and the parameters was greatly hindered by thermal effects in the transistors. In spite of the fact that the construction of the transistors ensured good cooling (the specimens were constructed similar to the regular P8 transistors, and the crystals were soldered to the crystal holders over the entire drain plane), the temperature did not stay constant enough during the measurements, owing to the strong localization of the heat relased in the field-effect transistors. Thus, the saturation current of the gate doubled in 20 minutes when $P_S = U_S I_S$ was 300 mw, and quadrupled for $P_S = 500$ mw (it then remained constant and returned to its previous value as the specimen was cooled): here $U_S = V_S - V_{SO}$.

The thermal effects resulted in the following: 1) appearance of negative conductivity in the gate at sufficiently high drain voltages; 2) reduction in the values of the saturation voltage and saturation current; 3) occasional appearance of negative drain — source conductivity in the saturation mode.

The first effect is attributed to the increase in the heat generated by the minority carriers at low gate voltages, when maximum power is generated at the drain. The two other effects are connected with the reduction in the mobility of the carriers upon heating. We therefore employed in all the measurements forced air cooling of the specimens, which reduced the heating considerably and suppressed the thermal effects.

Inasmuch as in our specimens φ_c was on the order of 1 — 2, a limitation of the drift velocity should have been observed. Figure 6 shows the calculated characteristic [formula (3)] for one of the specimens, along with the experimental points. The origin of the coordinates was shifted for each voltage between the gate and source by the amount of this voltage along the theoretical curve. As can be seen, the agreement is very good, with the exception of the high-current points at $V_{so} = 0$, where a limitation of the drift velocity should appear.

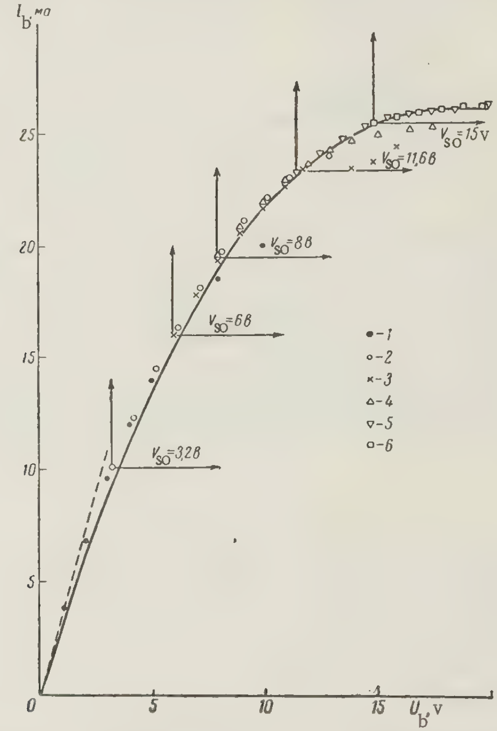


Figure 6. Calculated channel function ($V_0 = 18$ v, $I_0 = 65$ ma), and experimental points, plotted in shifting coordinates I_b, U_b (the bias voltage is equal to V_{so}):

1 — $v_{so} = 0$; 2 — $v_{so} = 3.2$ v; 3 — $v_{so} = 6$ v; 4 — $v_{so} = 8$v; 5 — $v_{so} = 11.6$ v; 6 — $v_{so} = 15$ v.

The maximum mutual conductance G_m was for this reason always less than the conductivity. In Table 1 are compared the values of G_m and G_0 for several specimens.

An asymmetry of the characteristics was observed for the majority of the specimens when the source and drain change places. Figure 7 shows a family of characteristics with strong asymmetry (specimen No. 365). The points at which the injection begins at $U_S > 0$ fit well on the volt-ampere characteristic $I_S = f(U_S)$ with $V_{SO} = 0$, when the source and the drain change places. The asymmetry of the characteristics is apparently due to the inhomogeneity of the channel (See Section **2**).

The breakdown voltage at various source voltages satisfied well the equation

$$(V_{dr})_e = (U_{dr} + V_{so})_e = V_e = \text{const.} \tag{18}$$

This is seen in Figure 7 (for both half-families), and also in Figure 8, which shows plots of $I_s = f(V_{so})$, for various U_s.

Table 1

No. of specimens	Maximum Current	Maximum mutual conductance v	Initial Conductivity	No. of specimens	Maximum Current	Maximum mutual conductance v	Initial Conductivity
593	20	3.2	4.2	503	20	4.0	4.75
595	30	2.4	4.5	311	42	6.4	6.7
808	20	1.8	2.45	545	40	6.4	—
				834	5.2	1.3	1.3

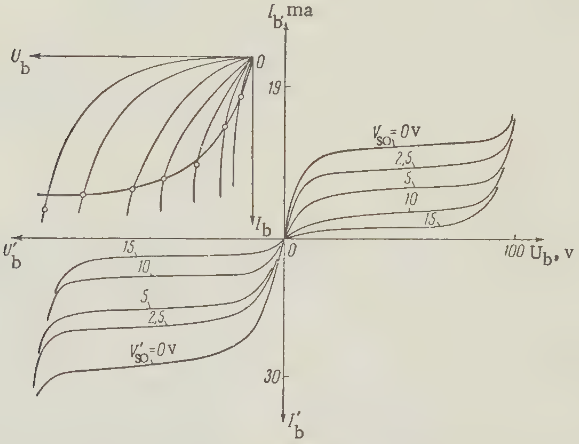


Figure 7. Volt-ampere characteristics $I_b = f(U_b)$ for the "forward" and "backward" transistor connection:
a — Normal family; b — family of characteristics when the source and drain change places: c — normal family for $U_b < 0$.

The internal resistance (differential resistance of the drain) of transistor specimens with maximum current 20 — 30 ma ranged from 10 to 100 kilohms, reaching a gentle maximum half-way between the saturation voltage and the breakdown voltage.

The input current ranged for most specimens from 10 to 30 microamperes. The input differential resistance in the case of sufficient cooling from the specimen was positive and its value for the good specimen ranged from 1 to 10 megohms.

The blocking ratio, i.e., the ratio of the saturation drain current at $V_{so} = 0$ to the minimum drain current (see Figure 8) exceeded 100 for the satisfactory specimens.

The capacitance introduced by the transistor into the "anode" circuit when the drain was biased by 10 — 15 v amounted to 20 — 40 micromicrofarads.

Preliminary measurements have shown that the transistors could amplify at frequencies on the order of several times ten Mc, and resonant circuits could be tuned by varying the voltages on the drain and on the gate.

2. Phototriode Properties of the Specimens

The possibility of using unipolar field-effect transistors as phototriodes was regarded in reference [5]. It is based on the high sensitivity, which can be obtained with germanium

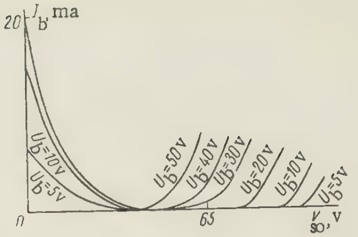


Figure 8. Volt-ampere characteristics $I_b = f(v_{so})$ for different values of U_b.

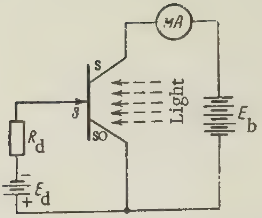


Figure 9. Diagram of phototriode connection of field-junction transistor.

and silicon phototriodes. In the case of a field-effect transistor, the p-n junction that separates the gate from the drain, channel, and source, can be used as a phototriode, while the voltage delivered to a load in the gate circuit will control the drain current (see Figure 9). Apparently the sensitivity of such a phototransistor has a value (for $R_{dr} \ll R_r$, where R_r is the differential resistance of the gate)

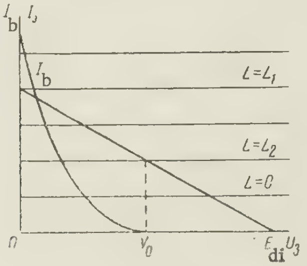


Figure 10. Used in the calculation of the transistor photosensitivity:

Volt-ampere characteristic $I_b = f(V_{so})$ of the transistor in the saturation mode and the characteristics of the gate as a photodiode.

$$M_{da} = M_{di} R_{di} G, \qquad (19)$$

where M_{di} is the sensitivity of a gate as a photodiode. When $R_{di} \gg R_i$ we have

$$M = M_{di} R_i G \qquad (20)$$

Taking $M_{di} \simeq 20$ ma/lumen, $R_i \simeq 10$ megohm, $G \simeq 3$ ma/v, we obtain $M_T = 600$ amp/lumen.

The dependence of the drain current on the illumination level, and also on the choice of the values of R_{di} and E_{di}, can be readily ascertained by using the photodiode characteristics of the gate and the drain-gate characteristic of the transistor (see Figure 10). We denote by L the light flux in lumens, incident on the light-sensitive surface of the transistor, and by I_{da} and I_f the dark and full gate currents, respectively. Then when

$$I_f = I_{da} + M_{\text{д}} L \geqslant \frac{E_{di}}{R_{di}}$$

the voltage on the gate does not exceed several hundred millivolts ($U_{di} = V_{so} \simeq 0$), and the channel current is close to the maximum value, I_{s0}. When

$$I_f = I_{da} + M_{di} L \leqslant \frac{E_{di} - U_{di\,bl}}{R_{di}}$$

the gate voltage U_{di} exceeds the blocking voltage $U_{di\ bl}$, and the channel current is close to zero. At light fluxes $L_1 \geqslant L \geqslant L_2$, where

$$L_1 = \left(\frac{E_{di}}{R_{di}} - I_{da}\right)\frac{1}{M_{di}}, \quad L_2 = \left(\frac{E_{di} - U_{di\,bl}}{R_{di}} - I_{da}\right)\frac{1}{M_{di}}, \qquad (21)$$

The drain current is refulated by the incident light. This is illustrated in Figure 11, which shows the experimental curves $I_s = f(L)$ for several values of E_{di}. By varying the value of E_{di}, we can obtain the intervals of sensitivity of the phototriode near any value of illumination, and the width of the interval (and consequently the sensitivity in it) is determined by the value of R_{di} (so long as $R_{di} < R_i$).

In plotting the relationships shown in Figure 11, and in other photoelectric measurements, we made use of a prefocused incandescent lamp, which produced under normal filament conditions a light flux of 1750 lumens (at a color temperature $\gtrsim 2,850$ degrees K). To calculate the sensitivity in amp/lumen we measured the area of the light-sensitive surface of the specimens, which in our case was the portion of the source surface not covered by the contact.

Table 2 lists the maximum sensitivities for four specimens. In measurements made on these specimens, the operating point was chosen in the region where the specimen had the greatest transconductance. We measured the increment in the drain current with the illumination varying by 50 — 120 lux.

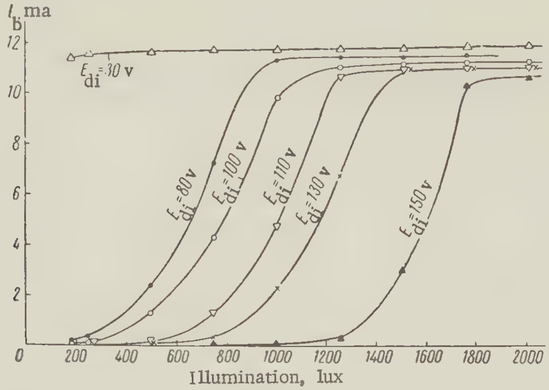


Figure 11. Lux-ampere characteristics of the drain current for different voltages E_{dr}(U_b = = 4 v; R_{di} = 1.9 megohms) .

Table 2

No. of Specimens	Area of light-sensitive surface, mm^2	Increase in illumination, lux	Increase in drain current, ma	Sensitivity, amp/lumen
346	2.0	56	4.9	44
380	0.5	50	5.4	216
365	1.0	120	8.6	71
1	0.6	59	5.65	159

3. Temperature Dependence

We investigated the influence of the temperature on the phototriode operation (Figure 9) . In a phototriode circuit, the effect of the temperature is quite analogous to the effect of additional illumination, as can be seen from the dependence of the drain current on the temperature at various values of E_{di}, as shown in Figure 12. The additional effect is that

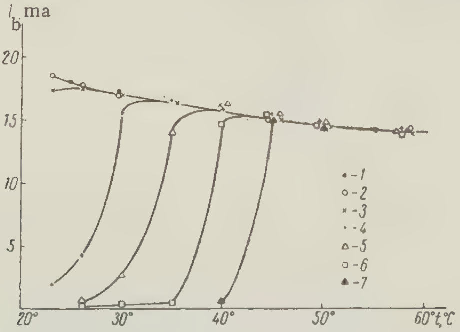


Figure 12. Dependence of the drain current on the temperature for different voltages. E_{di} (U_b = 8 v; R_{di} = 1.9 megohm) : 1 — E_{di} = 0 v; 2 — E_{di} = 20 v; 3 — E_{di} = 30 v; 4 — E_{di} = 50 v; 5 — E_{di} = 70 v; 6 — E_{di} = 100 v; 7 — E_{di} = 130 v.

with increasing temperature the maximum current decreases. This drop is due to the decrease in the electron mobility with increasing temperature, and its magnitude coincides with the value calculated on the basis of this assumption.

The sensitivity of the gate current to the temperature by the same factor GR_{di} (or GR_i) entering into the corresponding photosensitivities, and thus the situation is analogous to what takes place in ordinary phototriodes. It is, therefore, advantageous to use field-effect phototriodes as receivers for modulated radiation or for individual light pulses, when the resistance R_{di} in the circuit of Figure 9 can be replaced by an inductance of sufficient magnitude or, at constant modulation frequency, by a resonant circuit of high Q.*

When the gate bias source is sufficiently strong, such a circuit maintains high sensitivity to an ac light signal, but does not respond to constant illumination and to temperature variations in narrow ranges.

CONCLUSION

A possibility is indicated of obtaining field-effect transistors of the end type with high mutual conductance and considerable working currents at high blocking ratio. An investigation of such transistors has indicated the need for taking into account the saturation of the drift velocity of the electrons in strong electric fields during the analysis of the operation of the transistors. Field-effect transistors can serve as sensitive phototriodes. Their photosensitivity can exceed the sensitivity of other known semiconductor devices. The temperature dependence of the drain current at $T < 70$ degrees C is due to the temperature dependence of the mobility of the electrons and to the temperature dependence of the bias on the gate as a result of changes in the gate current. The former dependence is small, while the latter can be suppressed by using sources of constant gate bias with small output resistance, and by using inductances, resonant circuits, and other elements of low active resistance as control elements. When operating with large resistance in the gate circuit, the transistors can be used as highly sensitive temperature transducers in small temperature intervals.

The authors consider it their pleasant duty to thank V.A. Fomenko and K.M. Krolevets for a discussion of several problems considered in this paper.

REFERENCES

1. W. Shockley, Proc. I.R.E. 1952, 40, 11, 1365.
2. G.C. Dacey, I.M. Ross, Proc. I.R.E., 1953, 41, 8, 970.
3. G.C. Dacey, I.M. Ross, Bell System Techn. J., 1955, 34, 6, 1149.
4. N. Nelson, Proc. I.R.E. 1958, 26, 6, 1062.
5. J.B. Gunn, J. Electronics and Control, 1956, 2, 1, 87.

Submitted to the editors, 27 October, 1960

---

*This possibility was pointed out to us by V.S. Yanovich, who employed an analogous control circuit for ordinary phototriodes with controllable base, for which the authors are deeply grateful.

ON THE THEORY OF FIELD EMISSION AND THERMIONIC FIELD EMISSION OF METALS AND SEMICONDUCTORS

M. I.Yelinson, F.F. Dobryakova, V.F. Krapivin, Z.A. Malina and A.A. Yasnopol'skaya

Numerical calculations are given for the density of the field-emission and thermionic-field-emission currents, and also for the energy distribution of the emitted electrons for metals and semiconductors, using a new expression for the transparency of the surface barrier (3) and over a wide range of the parameters E, T, μ, and χ.

The results of the calculations differ from those previously obtained for metals on the basis of the WKB method.

Many data are obtained of importance for experimental investigations and operation of the cathodes.

The calculations for the semiconductors make it possible to estimate the limits of applicability of the simplest analytic expressions, and also to determine many parameters of the emitter substance.

1. STATUS OF THE THEORY

As is well known, the quantum-mechanical theory of field-emission of metals and semiconductors is based on several assumptions.

1. The "internal potential" of the metal or semiconductor is constant and agrees with the "bottom" of the conduction band, which is separated from the vacuum level by the magnitude of the electron affinity χ.

2. The potential barrier on the boundary with vacuum is a barrier of the image forces $V = -e^2/4x$. The total potential function is given by the curve ABx_0CG (Figure 1). The position of the point x_0 is determined from the equation $\chi = eEx_0 + e^2/4x_0$.

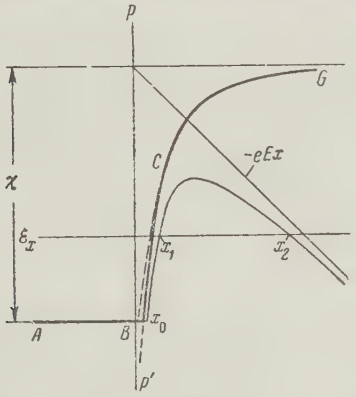


Figure 1. Potential barrier on the boundary between a solid and vacuum.

3. The flux of electrons incident on the emitting surface is calculated from the equilibrium distribution function over the states.

4. The problem is assumed to be one-dimensional; this is manifest in the dependence of the potential only on a single coordinate, neglecting the "spottiness" of the surface, the "atomic structure" of the surface barrier, and also in the calculation of the transparency of the barrier with the aid of the one-dimensional Schroedinger equation, where the transparency is determined by the energies connected with the normal components of the electron momentum, and is completely independent of the transverse components.

5. The equal energy surfaces are assumed spherical.

Assumption 1, corresponding to the Sommerfeld model, is not too well founded, but cannot, as is well known, be circumvented by simple means.

Assumption 2 is fairly correct starting with a certain distance from the surface (for example, from the point C, $x \simeq 3\text{Å}$, Figure 1). The variation of the potential function at closer distances has not yet been determined theoretically. Various heuristic assumptions were suggested in this regard quite some time ago (Langmuir, Schottky), but these are difficult to verify. Recently, in connection with experimentally observed anomalies in the volt-ampere characteristics of field emission at limitingly high currents, the hypothesis was again advanced [1] that the true barrier differs from the one obeying the image-force law (Curve ABCG, Figure 1). The anomalies should be particularly strong in the case of complicated barrier forms (presence of adsorbed layers). An attempt at a quantitative processing of the equations of emission under certain spatial assumptions on the barrier $V = -\frac{e^2}{4x + \frac{e^2}{\chi}}$ was made by G.F. Vasil'yev.

We note that the origin of the anomalies has not yet been established reliably. Nor has the correction to the image-force law for semiconductors been determined at the present time, and it is assumed in the theory that

$$V = -\frac{\varepsilon - 1}{\varepsilon + 1}\frac{e^2}{4x}$$

(ϵ is the dielectric constant).

Assumption 3 is valid for metals and low-resistance semiconductors, and is very poorly satisfied for high-resistance semiconductors with low values of χ. [3].

Assumptions 4 and 5 have little justification, but, as in case 1, they cannot be circumvented by simple means.

In most investigations the transparency of the potential barrier D is calculated with the aid of a quasi-classical (WKB) method. For the image-force barrier

$$D = \frac{|\Psi_{tr}|^2}{|\Psi_{inc}|^2} = \exp\left\{-\frac{4\sqrt{2}}{3}\frac{m^{1/2}|\mathscr{E}_x|^{3/2}}{\hbar eE}\Theta(y)\right\}, \tag{1}$$

where $\mathscr{E}_x = \frac{p_x^2}{2m} + V$; E is the external field and $\Theta(y)$ is the Nordheim function:

$$\Theta(y) = \frac{1}{\sqrt{2}}\sqrt{1 + \sqrt{1 - y^2}}\,[L(k') - (1 - \sqrt{1 - y^2})K(k')]. \tag{2}$$

In expression (2), $y = e\sqrt{eE}/|\mathscr{E}_x|$ (for metals), $k' = (2\sqrt{1 - y^2}/1 + \sqrt{1 - y^2})^{1/2}$; L(k') and K(k') are the complete elliptic integrals of the first and second kinds.

As is well known, a shortcoming of the WKB method is the indeterminacy of the solutions for the Ψ functions in the classical "turning points" x_1 and x_2 (see Figure 1) and the fact that this method is poorly applicable for non-smooth barriers and for barriers having a small area under a given level $\mathscr{E}_x$.

Thus, the method becomes unreliable in principle in the calculation of thermionic field emission of substances with low work functions. Yet it is widely used in analytic [4] and numerical [5] calculations for metals, and also in analytic calculations for semiconductors [6].

In order to construct a general theory of thermionic field emission and field emission [7], use has been made recently of an approximate expression for the transparency of the potential barrier, suitable in principle for electrons of arbitrary energy, both below and above the peak of the barrier [8]. For the case of the image-force law we have

$$D = \left\{1 + \exp\left[\frac{4\sqrt{2}}{3}\frac{m^{1/2}|\mathscr{E}_x|^{3/2}}{\hbar eE}\Theta(y)\right]\right\}^{-1}. \tag{3}$$

However, in view of the complexity of the problem, numerous approximations have been used for the calculation of the current, and these have led to narrow temperature and field intervals in which the final analytic expressions are valid. Certain approximations signify essentially a return to the WKB method. In addition, the formulas are cumbersome and inconvenient for use of experimenters. This remark applies equally well to metals and to semiconductors, although certain formulas for semiconductors (the case of nondegenerate gas) have wider ranges of validity [9].

This status of the theory has motivated us to carry out numerical evaluation of the density of the field emission and thermionic field emission currents and also of the energy distribution in metals and semiconductors, using expression (3). The purpose is to extract the maximum amount of information necessary for comparison of experiment with theory and to estimate the range of validity of the simplest approximate analytic expressions. We have decided to retai the equilibrium distribution function of the electron states, inasmuch as it is valid for a large number of substances. This preserves the possibility of comparing our calculations with those of references [5] and [6]. For the same reason, we decided to retain the image-force barrier, which, furthermore, is physically just as well founded as are other reasonable barrier forms.

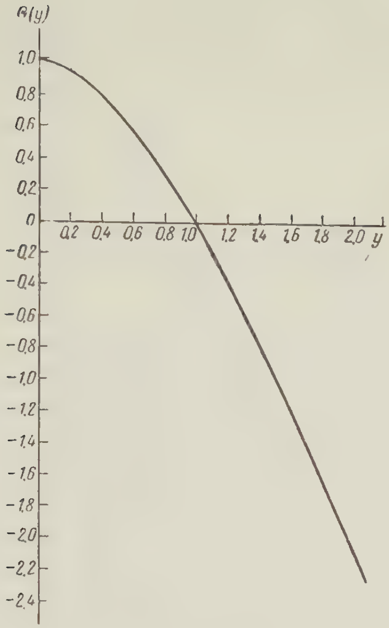


Figure 2. Plot of the Nordheim function Θ(y).

2. METALS

Our calculation of the density of the field-emission and thermionic-field-emission currents for metals is based on the integral

$$j = \frac{4\pi mekT}{h^3} \int\limits_{-\infty}^{+\infty} \frac{\ln\left\{1 + \exp\left[-\frac{\mathscr{E}_x - \mu}{kT}\right]\right\} d\mathscr{E}_x}{1 + \exp\left\{\frac{4\sqrt{2}}{3} \frac{m^{1/2}|\mathscr{E}_x|^{3/2}}{\hbar eE} \Theta\left(\frac{e\sqrt{eE}}{|\mathscr{E}_x|}\right)\right\}}. \quad (4)$$

The integrand is the product of the equilibrium current of degenerate-gas particles with energy $\mathscr{E}_x$ in the interval $\mathscr{E}_x, \mathscr{E}_x + d\mathscr{E}_x$, multiplied by expression (3). The energy is referred to the vacuum level. The quantity $-\chi$ is replaced by $-\infty$, in view of the large energy interval of the filled part of the conduction band.

The integral is on the whole complicated, but a particular complication is introduced into the calculations by the Nordheim function Θ(y). In the WKB method it is sufficient to know the function Θ(y) in the interval $0 \leqslant y \leqslant 1$, for which it was recently tabulated [10] For our calculations we had to tabulate Θ(y) for the region $y \geqslant 1$, with the aid of the general expression (2). A plot of Θ(y) is shown in Figure 2. In the interval $0 \leqslant y \leqslant 1$ a good agreement is noted with the data of [10]. In the choice of the parameters E, T, and μ we attempted to include among them those used by the authors of reference [5], who used expression (1) in their calculations.

The results of the calculations are shown in Figure 3, in the customary coordinates $\log j = f(1/E)$ for the interval $\varphi = 2 - 4.5$ ev and for $T = 0 - 3000$°K.

The general course of the voltage-ampere characteristics is best traced for material with low work function $\varphi = -\mu = 2$ ev. Figure 4 shows curves of $\log j = f(1/E)$ (Curve 1) and $\log(j/E^2) = f(1/E)$ (Curve 2) for $kT = 0$ and $\varphi = 2$ev. At small j, curve 1 is a straight line. Further at a field E much smaller than the critical field E_{cr}, for which $y = e\sqrt{eE_{cr}}/\varphi = 1$ the curve departs from a straight line towards lower values of j. A noticeable tendency is seen towards a faster increase in j at fields E greater than critical. At still greater fields, the presence of Schottky deviations can be assumed.

Curve 2 reflects most prominently the fact that the deviation from linearity begins long before E_{cr} is reached. A noticeable difference from the calculated results of reference [5], in which the WKB approximation was used, is evident. The difference can be naturally explained by the fact that at small E we have $\left\{\frac{4\sqrt{2}}{3} \frac{m^{1/2}|\mathscr{E}_x|^{3/2}}{\hbar eE} \Theta(y)\right\} \gg 1$, and both results should coincide. As E increases, the exponential function comes close to unity, and the new transparency of the barrier is much lower than the quasi-classical one. The typical form of the experimental characteristics $\log j = f(1/E)$ is similar to the form of Curve 1 on Figure 4. In many cases however, the departure from linearity towards decreasing j and the subsequent faster increase in j occur at electric fields which are apparently even smaller than E_{cr}.

An interesting situation is brought about by the fact that a similar type of characteristic can be obtained also by assuming a barrier form different from the image-force type, or by taking into account the space charge connected with the emission.

The results of the calculation permit a theoretical investigation of the energy distribution of the emitted electrons for E and T varying over a wide range. In the case of field emission, the distribution over the total energies $P(\mathscr{E})d\mathscr{E}$ differs noticeably from the distribution over the normal components $P(\mathscr{E}_x)d\mathscr{E}_x$. There is no calculation of $P(\mathscr{E})d\mathscr{E}$ in [5]. We do not give here all the calculated distributions, Figure 5 shows only several distributions of $P(\mathscr{E})$ in relative units. the zero corresponds to the level μ; $\varphi = 4.5$ ev (for comparison with Figure 6).

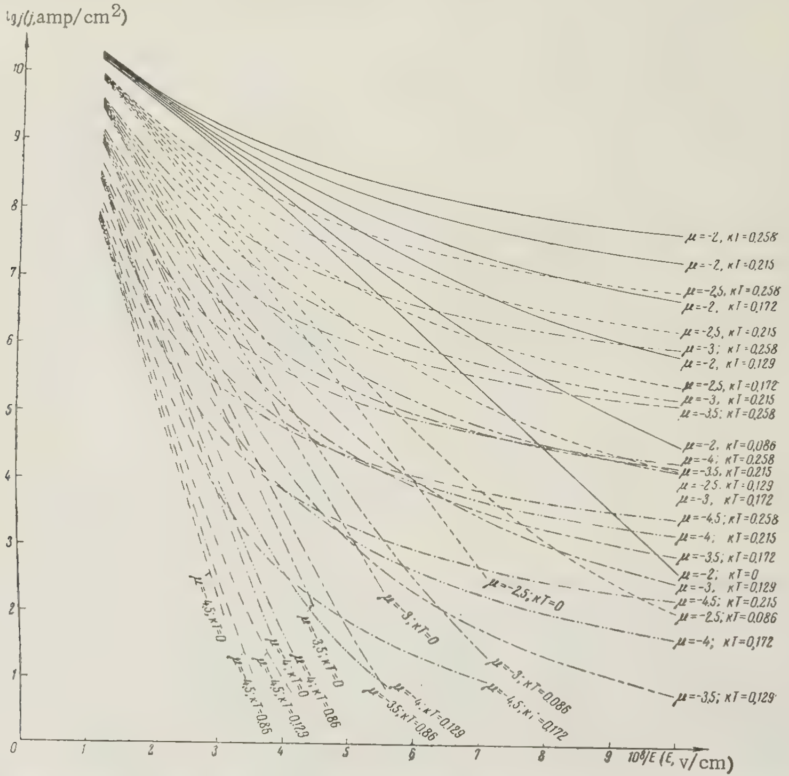


Figure 3. Dependence of log j = f(1/E) obtained by numerical calculation for metals using Equation (4) (for the interval $\varphi = 2 - 4.5$ ev and T = 0 — 3000 degrees K).

When E = const, the maximum shifts towards lower energies with increasing E and the distribution first becomes narrower, and then remains practically constant.

The distributions $P(\mathscr{E})$ determine the noise properties of the cathodes. It is found that that the half-width of the distribution over the total energies is always less than the correspondin half-width for the energies $\mathscr{E}_x$. However, precisely the latter were used for an estimate of the effective temperature of autoelectronic cathodes [11] because the effective temperature was found to be greatly overestimated.

Figure 6 shows several distributions for the energy $\mathscr{E}_x$ (solid curves). The transformation of the form and of the width of the curves for constant T and variable E, as well as for constant E and variable T, is quite obvious. The dotted curves represent the corresponding results of the calculation [5]. It is patently evident that the use of the WKB approximation exaggerates the role of thermally excited electrons in the emission.

The calculation results permit a choice of various operating modes for emitters, i.e., an optimal choice of E and T. Thus, for example, if with increasing T it is possible to maintain j by reducing the field E from a value E_1 to E_2, then it is possible to increase the total current, for the same voltage U, by a factor $(E_1/E_2)^2$.

Figure 7 shows a series of equal-current curves in coordinates E, T, for $\varphi = 2.5$, 3.5 and 4.5 ev, and for four values of j. It is seen from the curves that it is difficult to reduce E greatly by increasing T. It must be noted, however, that at very high voltages, any possibility of reducing the voltage is important.

From the curves of Figure 3, we can extract important information on the variation of the slope of the characteristics at different modes, $dj/dE = \gamma\, di/dU$, and also concerning

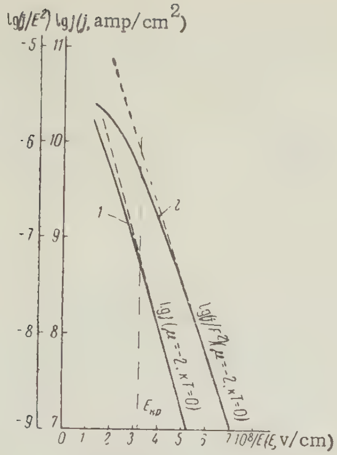


Figure 4. Typical plots of log j = = f(1/E) and log (j/E^2) = f(1/E) for metals.

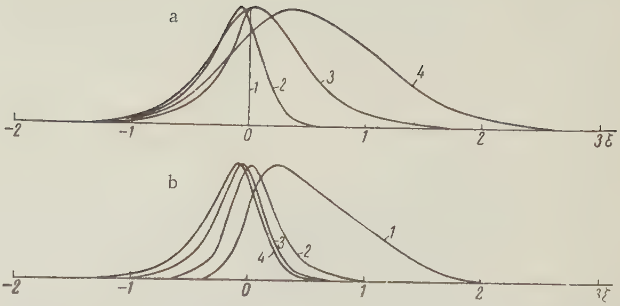


Figure 5. Distribution of emitted electrons over the total energies for metals.
a — E = 7.2 · 10^7 v/cm, μ = -4.5, 1 — kT = 0, 2 — kT = 0.086, 3 — kT = 0.172, 4 — kT = 0.258; b — kT = 0.086, μ = -4.5, 1 — E = 2 · 10^7 v/cm, 2 — E = = 3.4 · 10^7 v/cm 3 — E = 5.2 · 10^7 v/cm 4 — E = 7.2 · 10^7 v/cm.

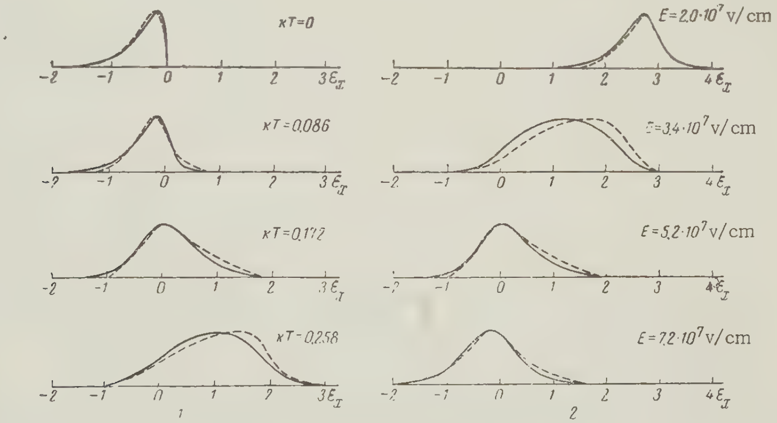


Figure 6. Distribution of emitted electrons over the energies $\mathscr{E}_x$ for metals: 1 — E = 5.2 · · 10^7 v/cm, μ = 4.5 ev, 2 — kT = 0.172, μ = - 4.5 ev.

their nonlinear properties, connected with the value of d^2j/dE^2, something which we shall not discuss here.

The results of the calculations allow us to propose a simple method of determining the work function φ of the emitter substance and the surface field E, without calculating the latter with the aid of complicated and unreliable electrostatic formulas.

To determine φ and E, it is necessary to use in principle the experimental volt-ampere characteristics log i = f(1/U), plotted for two different emitter temperatures. It is assumed that the temperature emitter is measured with sufficient accuracy and that its emitting surface is well known from electron photographs. In the determination of the quantities indicated above, certain errors creep in, but it is well known that the possible errors are much lower in this case than those connected with the calculation of the field E.

The method proposed is illustrated in Figure 8. The curves a and b are the theoretically calculated relatioships log (j_{2000}/j_0) = f(φ) and log (j_{2000}/j_{1000}) = f(φ), where φ is the work function and j_{2000}, j_{1000}, and j_0 are the current densities at the corresponding temperatures. The curves on these plots are constructed under constant values of j_0 or j_{1000}. Curve c shows schematically the experimental characteristics $\log j = f(1/U)$ (j is determined from the emitting area) .

By specifying j_0 or j_{1000}, say $j_0 = 10^6$ am/cm^2, we can determine from these characteristics the value of j_{2000}, and then determine φ from the theoretical curves. From the value of φ we can determine in the final analysis the value of $d(\log j)/d(1/E)$. Knowing the experimental value of $d(\log j)/d(1/U)$, we can determine the geometric factor $\beta = E/U$, and consequently also the field on the emitter vertex (E) . We can thereupon determine experimentally, with sufficient accuracy, the range of operating electric fields, a very important factor for the physical interpretation of the experimental results.

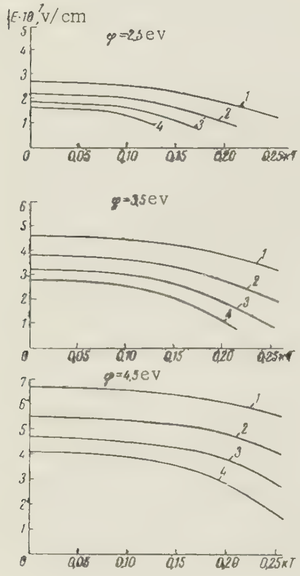


Figure 7

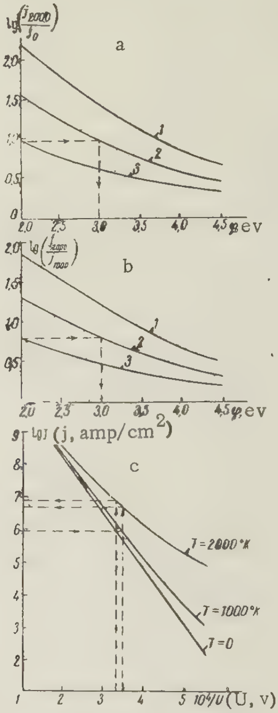


Figure 8

Figure 7. Equal current-density curves for simultaneous variation of E and T, and values $\varphi = 2.5$, 3.5, and 4.5 ev. 1 — $j = 1 \cdot 10^7$ amp/cm^2; 2 — $j_0 = 10^5$ amp/cm^2, 3 — $j = 1 \cdot 10^6$ amp/cm^2 4 — $j = 1 \cdot 10^4$ amp/cm^2.

Figure 8. Illustration of the method used to determine the surface electric field E and the work function φ; a — Theoretical relationship: 1 — $j_0 = 10^5$ amp/cm^2, etc. b — Theoretical relationship: etc. c — Experimentally plotted relationship.

The method described here is similar to the well-known method of I.S. Andreyev [12], but does not require any cumbersome calculations and is not based on the analytical formula of Guth and Mullin, the range of application of which is limited.

It is important to note that if the initial values chosen for j_0 and j_{1000} are not too large, the results of the determination of φ and E will be valid even when the true surface barrier deviate from the image-force law.

Thus, the calculations for metals enable us to obtain many interesting consequences, impo tant both for the understanding of the mechanism of emission, and for experimental research.

3. SEMICONDUCTORS

We use, as the basis for the calculation of the density of field-emission and thermionic-field-emission currents for semiconductors, the integral

$$j = e\int_0^\infty n\left(\frac{1}{2\pi mkT}\right)^{1/2}\exp\left(-\frac{\mathscr{E}_x}{kT}\right)\frac{d\mathscr{E}_x}{1+\exp\left\{\frac{4\sqrt{2}}{3}\,\frac{m^{1/2}|\chi-\mathscr{E}_x|^{3/2}}{\hbar eE}\,\theta(y)\right\}}. \tag{5}$$

For convenience, the energy is referred to the bottom of the conduction band. In addition, as is readily shown, we have in this case

$$y = \left(\frac{\varepsilon-1}{\varepsilon+1}\right)^{1/2}\frac{e\sqrt{eE}}{|\chi-\mathscr{E}_x|}.$$

The integral (5) is based on the notion of nondegenerate electron gas over the entire interval of fields E. We consider this case to be the most characteristic and most interesting for field emission of semiconductors.

The electron gas in the semiconductor can be degenerate, independent of the application of the external field. In the case the formulas for metals yield highly accurate values of the density of the field-emission current. One can expect degeneracy of electron gas in the semiconductor layers next to the surface, under the influence of the penetration of the external field, only in the case of substances having very high χ, when there is practically no current drawn at high fields to disturb the equilibrium character of the processes.

In cases of practical interest, namely those of substances with relatively high resistance and low χ, electric breakdown and destruction of the emitter precede the degeneracy.

Inasmuch as n depends on T in the nondegenerate case, it is advantageous to calculate j/n. To change over to j we must carry out a parallel displacement of the characteristics $\log(j/n) = f(1/E)$ along the ordinate axis.

Certain difficulties in the choice of parameters for the numerical integration are caused by the relation between the electron affinity χ and the dielectric constant ϵ. The quantity ϵ is connected with χ (for example, assuming the image-force law, by the relation $\chi = \frac{e^2}{2a}\frac{\epsilon-1}{\epsilon+1}$, where a is the lattice constant). At the same time, ϵ enters into the formula explicitly via the argument of the Nordheim function.

We deemed it possible to specify χ and ϵ independently (in particular, different values of χ for identical ϵ), for this ensured a large set of possible combinations. In principle, it is always possible to choose suitable semiconductors.

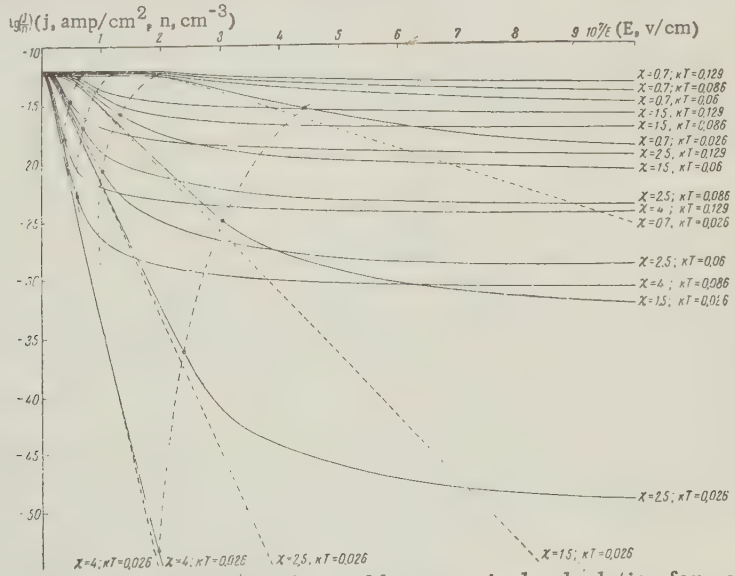


Figure 9. Plot of log (j, n) = f(1/E) obtained by numerical calculation for semiconductors, using Equation (5), for the interval χ = 0.7 — 4 ev, T = 300 — 1500 °K., and ϵ = 6; dotted curves correspond to Stratton's formula (7) for T = 300 °K. Dash-dot lines are plotted on the basis of criterion (6).

An integral analogous to (5) (with WKB transparency) was calculated in reference [6] by an excessively simplified method. Subject to the condition

$$eE \gg \frac{2\sqrt{2}\, m^{1/2}}{\hbar} \chi^{1/2} kT \tag{6}$$

(Stratton's criterion) the transparency is practically independent of the electron energy and corresponds, for example, to the transparency for the bottom of the conduction band.

In this case we obtain a simple analytic formula for j:

$$j = ne\left(\frac{kT}{2\pi m}\right)^{1/2} \exp\left\{-\frac{4\sqrt{2}}{3}\,\frac{m^{1/2}\chi}{\hbar eE}\,\Theta(y)\right\}. \tag{7}$$

The main problem in the calculation is to obtain the exact theoretical variation of the volt-ampere characteristics by using a new expression for the transparency, and a comparison of these characteristics essentially with the simpler formula (7).

Figure 9 shows the characteristics log (j/n) = f(1/E) calculated for the interval of values χ = 0.7 — 4 ev and kT = 0.026 — 0.029 ev, and for a typical value $\epsilon = 6$ (we omit the analogous series of curves for other values of ϵ).

It is easy to note several interesting features in these characteristics. First, the curves exhibit saturation at values E_{cr} which are defined for each χ. This fact distinguishes the behavior of semiconductors from that of metals, for which a complicated variation of the curves is observed after E_{cr} is reached (Curves 1 of Figure 4).

It is still difficult to propose at present a rational explanation of this difference.

The calculated values of E_r agree well with the simple formula $\frac{\epsilon - 1}{\epsilon + 1}^{1/2} \frac{e\sqrt{eE_{cr}}}{\chi} = 1$ and this result is by no means trivial.

We note, first, that the limiting value of j/n increases with increasing T. This is readily explained by the fact that at a given concentration, the particle current incident on the surface increases with increasing average thermal velocity.

Second, the slope of the quasi-linear portions of the characteristics decreases noticeably with increasing temperature. This is particularly noticeable for small χ, but manifests itself quite cleanly also for large χ.

Third, the interval j is covered by the quasi-linear portions of the characteristic decreases sharply with increasing temperature.

Fourth, at relatively low values of E, thermoelectronic "tails" are clearly evident.

The dotted curves in Figure 9 are the characteristics log (j/n) = f (1/E), plotted in accordance with Equation (7) for kT = 0.026 ev (room temperature). The characteristics for higher temperatures are obtained by parallel shift along the ordinate axis. The dash-dot lines of Figure 9 delineate for each temperature two regions, with Equation (6) applicable on the left-hand region (for larger E).

A comparison of the exact characteristics with expression (7) leads to the following conclusions.

1. The validity of Stratton's theory is greater the lower T and the higher χ. With increasing temperature, the slopes of the Stratton lines and of the quasi-linear portions of the true characteristics become noticeably different. At low χ and increased T, there is almost no region of E where the Stratton theory is applicable.

2. The minimum value of E (at kT = 0.026 ev) in accordance with (6) practically corresponds to the end of the quasi-linear portion of the true characteristic. The strong inequality sign in (6) is excessive, and at low temperatures Stratton's theory is valid over a wider range of E.

3. At low temperatures, the slope of the quasi-linear portions of the characteristics tan α is close to the slope of the Stratton lines tan α_c.

Figure 10 shows a plot of σ = tan αc / tan α as a function of the affinity χ for kT = 0.26 ev.

Figure 10 leads to a conclusion of importance to the experimenter, namely that the electron affinity can be determined from the experimental characteristics, by using the simple formula

$$\tan\alpha = \frac{d(\lg j)}{d\left(\frac{1}{E}\right)} = -6.85\cdot 10^{7}\chi^{3/2} S(y), \tag{8}$$

where

$$S(y) = \Theta(y) - \frac{y}{2}\frac{d\Theta}{dy}.$$

The same figure shows the dependence on χ of the number of orders of j/n, included in the quasi-linear region at $kT = 0.026$ ev; this dependence is of great significance for experimental research.

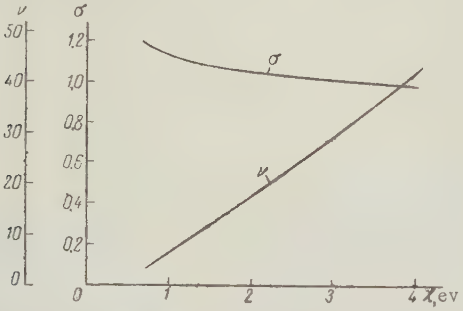


Figure 10. Plot of $\sigma = \tan\alpha_c/\tan\alpha$ and the number of orders ν of the quasi-linear region of characteristics $\log(j/n) = f(1/E)$, as functions of χ.

Measurements of field emission, as is clear from the foregoing, make it possible at the same time to determine experimentally such important semiconductor parameters as χ, ϵ, n, and also the conduction activation energy, using parallel shift of the saturation region. Thus, calculations for semiconductors lead to important deductions of value both for the understanding of the emission mechanism and for experimental research.

In the calculation of the emission-current density we purposely made no mention of effects typical of semiconductors, namely the penetration of the external field and the influence of the surface states of the electrons.

A correct account of these effects is usually very complicated. The methods developed in [6] are in many respects debatable. Thus, for example, calculation of the penetration of the field is based on electrostatic considerations at a constant level of the chemical potential, independently of the density of the current drawn. Yet it is obvious that this is utterly incorrect when $E/\epsilon \simeq j\rho$ (ρ is the specific resistivity of the semiconductor). i.e., in most cases of interest.

To calculate the near-surface concentration of the electrons it is necessary to employ other methods, such as those developed, for example, in reference [13]. Consequently, the frequently used expression $n = n_\infty \exp(\Delta R/kT)$, where n is the near-surface concentration, n_∞ the unperturbed concentration, and ΔR is the bending of the bands near the surface, is valid only for very small values of j. An account of the role of the surface states, given in reference [6], is correct only for certain special cases (two local systems of surface levels, dynamic exchange of electrons with the conduction band only).

It can be assumed that for semiconductors with large ρ and small χ, such as many ionic crystals are, it is more correct to disregard the penetration of the external field. With this the internal field in the semiconductor is determined by the passing current.

For such semiconductors we can neglect also the role of surface states, inasmuch as usually the work occurs at electric fields for which the barrier due to the surface states is "breached."

Many experimental characteristics [14] actually correspond to Figure 9.

CONCLUSION

1. The paper contains numerical calculations of the density of field-emission and thermionic-field emission currents, and also of the distribution of the emitted electrons by energies, for metals and semiconductors, using a new expression for the transparency of the surface barrier (3) and for a wide range of the parameters E, T, μ, and χ.
2. It is established that the volt-ampere characteristics differ from those calculated by the WKB method.
3. It is established that the WKB method leads to an overestimate of the role of thermally excited electrons in emission.
4. At low temperatures, the half-width of the distribution of the total energies is less than the distribution curve for $\mathscr{E}_x$.
5. A series of equal-current curves was obtained (under simultaneous variation of E and T), of importance in the operation of the cathodes.
6. A simple method is proposed for determining the work function and the surface electric field without calculating the latter, an important factor in experimental research.

7. It is established that the general course of the volt-ampere characteristics of emission from semiconductors differs appreciably from the course for metals at high current densities. It is shown that by investigating field emission it is possible to determine several semiconducto parameters.

8. The limits of applicability of the simple theory of reference [6] are examined. It is established that at high temperatures this theory if practically inapplicable, whereas at room temperature its validity limits are sufficiently broad and the theory can be reliably used in experimental research.

REFERENCES

1. T.J. Lewis, Phys. Rev. 1956, 101, 6, 1694.
2. G.F. Vasil'yev, Radiotekhnika i elektronika, 1960, 5, 11, 1857.
3. M.I. Yelinson, Radiotekhnika i elektronika, 1959, 4, 1, 140.
4. E. Guth, C.J. Mullin, Phys. Rev., 1942, 61, 4, 339.
5. W.W. Dolan, W.P. Dyke, Phys. Rev., 1954, 95, 2, 327.
6. R. Stratton, Proc. Phys. Soc., B, 1955, 68, 430, 746.
7. E.L. Murphy, K.H. Good, Phys, Rev., 1956, 102, 6, 1464.
8. S.C. Miller, K.H. Good, Phys Rev., 1953, 91, 1, 174.
9. G.F. Vasil'yev, Radiotekhnika i elektronika, 1958, 3, 7, 962.
10. R.E. Burgess, H. Kroemer, J.M. Houston, Phys, Rev., 1953, 90, 4, 515.
11. R.W. DeGrass, G. Wade, Proc. I.R.E., 1956, 44, 8, 1048.
12. I.S. Andreyev, ZhTF, 1953, 23, 5, 849.
13. K.B. Tolpygo, Radiotekhnika i elektronika, 1958, 3, 8, 980.
14. M.I. Yelinson, A.G. Zhdan, G.V. Vasil'yev, Radiotekhnika i elektronika, 1960, 5, 12, 20

Submitted to the editor, 21 November, 1960.

THERMAL RADIATION PRODUCED BY HOT CARRIERS IN SEMICONDUCTORS

Sh.M. Kogan

A theoretical calculation is made of the thermal radiation produced by hot carriers in a homogeneous semiconductor to which a strong homogeneous electric field is applied. Estimates of the resultant expressions for the radiation intensity show that this radiation can be observed experimentally. Certain features and possible application of the radiation are briefly discussed.

The presence of free carriers (conduction electrons and holes) in semiconductors bring about, as is well known, absorption of electromagnetic radiation (see, for example. [1]). Hence, it follows directly that the free carriers contribute to the thermal radiation of the semiconductor. The radiation produced by the carriers, like the absorption of electromagnetic radiation produced by these carriers, arises when the carriers are scattered by lattice vibrations and by impurities; this radiation is essentially bremsstrahlung.

An appreciable thermal radiation by the carriers can be expected, naturally, when the average energy $\bar{E}$ of the carrier random motion corresponds to sufficiently high temperatures. This takes place when a strong field is applied to the semiconductor, so that the carriers become "hot": their average energy becomes $\bar{E} \gg \varkappa T/2$, where T is the temperature of the semiconductor lattice and $\varkappa$ is Boltzmann's constant [2]. Thermal radiation of hot carriers can be observed when the electron (or hole) gas in a semiconductor is heated by a strong field.

To estimate the magnitude of the effect, let us calculate the spectral distribution of the radiation intensity of hot carriers in a homopolar semiconductor such as germanium. We make several assumptions to simplify the calculation. We assume that the mean-free-path l for momentum scattering is determined by the interaction between the carriers and acoustic phonons. This assumption does not exclude from consideration the case when the mean free path for the scattering of the carrier energy is greater than l, and is determined by interaction with optical phonons. The latter influences the form of the symmetrical distribution function $n_s(p)$, i.e., the energy distribution of the carriers E_p, and thereby the intensity of radiation.

We confine ourselves to a region of radiation frequencies ω satisfying the inequality

$$\omega \gg 1/\tau, \tag{1}$$

where τ is the effective mean-free-time of the carriers. In this case we can confine ourselves to an analysis of single-phonon scattering, since the electromagnetic field executes a large number of oscillations during the time τ between two successive scattering events.

Under conditions when inequality (1) is satisfied, the dimensions of the semiconductor specimen usually exceed the radiation wavelength, and consequently diffraction effects can be disregarded.

We shall assume that the energy absorption coefficient K is so small in the spectral interval of interest to us, that absorption of radiation inside the specimen can be neglected; the radiation is then simply proportional to the volume V of the specimen. This case is usually realized in the wavelength region beyond the edge of the intrinsic absorption of the semiconductor.

There is no point in going through the details of the calculation of the probability of radiation of a photon by an electron (or hole) through scattering on acoustic lattice vibrations, for it is sufficient to use the expression for the probability of photon absorption (see, for example, [1]). We obtain for the probability of radiation of a photon with wave number k and polarization in the direction of the unit vector e_λ:

$$w_{\text{emit}}^{\text{abs}} = \frac{(2\pi)^2}{V^2} \frac{e^2}{\varepsilon\hbar\omega^3 m^2} \frac{E_1^2\hbar\mathfrak{f}}{2\rho u} \left\{ \begin{matrix} \overline{N}_{\mathfrak{f}} \\ \overline{N}_{\mathfrak{f}} + 1 \end{matrix} \right\} (\mathfrak{f}e_\lambda)\, \delta\left(\omega_k \mp \Omega_{\mathfrak{f}} - \frac{E_p - E_{p\pm\mathfrak{f}}}{\hbar} \right). \tag{2}$$

The upper symbols refer to phonon absorption, and the lower to the process where a phonon is emitted upon radiation of the photon; f and $\hbar\Omega_f$ are the wave vector and energy of the phonon; $\overline{N}_f$ is the average number of phonons with wave vector f; ρ and ϵ are the density and dielectric constant of the crystal; $\hbar p$ and m are the quasi-momentum and effective mass of the electron (or hole);

$$E_p = \hbar^2 p^2 / 2m; \tag{3}$$

E_1 is the deformation-potential constant and u is the velocity of sound.

In the case of acoustic phonons

$$\overline{N}_f \approx \varkappa T / \hbar\Omega_f \gg 1.$$

In addition, we can neglect the energy of the phonon compared with the energy of the electron (or hole) and the energy of the photon ω.

The probability (2) must be summed over all final states (i.e., must be integrated over all f) and averaged over all initial states with the aid of the symmetrical part of the distribution function $n_s(p)$ (the antisymmetrical part does not contribute to the effect). After making the substitution (see [3])

$$\frac{\varkappa T E_1^2 m^2}{\rho u^2 \pi \hbar^4} = \frac{1}{v_p \tau_p} = \frac{1}{l}, \tag{4}$$

where v_p and τ_p are the velocity of carrier with momentum $\hbar p$ and its mean-free-time, respectively, we obtain the following expression for the spectral distribution of the radiation intensity within the full solid angle:

$$\frac{dI}{d\omega} = V \frac{4\sqrt{2} n e^2 \varepsilon^{1/2}}{3\pi m^{3/2} c^3 l} (\hbar\omega)^{3/2} \frac{1}{n} \int\limits_{\sqrt{2m\omega/\hbar}}^{\infty} \frac{2\cdot 4\pi}{(2\pi)^3} dp\, p^2 n_c(p) \left(\frac{2E_p}{\hbar\omega} - 1\right) \left(\frac{E_p}{\hbar\omega} - 1\right)^{1/2}. \tag{5}$$

In Equation (5), n is the number of carriers per unit volume and c is the velocity of light.

In that important particular case when $n_s(p)$ can be approximated by a Maxwellian function with some electron temperature T_e, i.e., when

$$n_c(p) = \frac{n}{2} \left(\frac{2\pi\hbar^2}{m\varkappa T_e}\right) \exp(-E_p/\varkappa T_e), \tag{6}$$

the expression for $dI/d\omega$ assumes the form

$$\frac{dI}{d\omega} = V \frac{32\sqrt{2} n e^2 \varepsilon^{1/2}}{3\pi^{3/2} l} \left(\frac{\bar{v}}{c}\right)^3 A\left(\frac{\hbar\omega}{\varkappa T_e}\right). \tag{7}$$

Here

$$\bar{v} = (\varkappa T_e / m)^{1/2}; \tag{7a}$$

$$A(z) = \frac{1}{2} e^{-z} \int\limits_0^{\infty} dx\, x^{1/2} (x+z)^{1/2} \left(x + \frac{z}{2}\right) e^{-x}. \tag{7b}$$

The function A(z) can be expressed in terms of Whittaker functions.

The electron temperature T_e as a function of the heating field F can be estimated in the following fashion. B.I. Davydov [4] derived for the case when the carriers interact only with acoustic phonons the following expression for T_e/T (with $T_e/T \gg 1$):

$$T_e/T = \frac{1}{2\sqrt{3}} (\sqrt{\varkappa T / m}/u)(eFl/\varkappa T). \tag{8}$$

T_e can be estimated in a different manner, too. In the case when the carrier scattering is characterized by a mean free path that is independent of the energy E_p, T_e is related to the mobility μ_F in a field by the following simple formula [5]:

$$T_e / T = (\mu_0 / \mu_F)^2, \tag{9}$$

where μ_0 is the mobility in weak fields. For n-type germanium at T = 77°K and F = 6,000 v/cm ($m \simeq 0.2 \times 10^{-27}$ g, $u \simeq 5 \times 10^5$ cm/sec, $l \simeq 5 \times 10^{-5}$ cm, μ 2×10^4 cm^2/v-sec, $\mu_F \simeq 1.5 \times 10^3$ cm^2/v-sec [6]), we form either formula (8) or formula (9)

$$T_e / T \sim 180.$$

From this we can conclude that in n-type germanium, at fields on the order of several kv/cm (prior to breakdown), we can obtain electron temperatures on the order of 10^4 K.

Let us calculate the intensity of thermal radiation of hot electrons in n-type germanium for $T_e = 10^4$ °K, $l = 5 \times 10^{-5}$ cm (T = 77°K), $n = 10^9$ cm^{-3}, and $V = 1$ cm^3.

We assume that the radiation detector registers a spectral region from $\hbar\omega = 0.15$ ev to $\hbar\omega = 0.30$ ev ($\Delta\ \ \omega = 0.15$ ev) in a solid angle equal to 1/4 of total. Then

$$\Delta I \sim 4 \cdot 10^{-6} \quad \text{watt*}.$$

This result enables us to conclude that it is feasible to observe experimentally the thermal radiation of hot electrons.

An experimental investigation of thermal radiation of hot conduction electrons or holes can yield information on the extent to which they are heated in a strong field, and particularly on their effective temperature.

An interesting feature of radiation of hot carriers is the possibility of modulating this radiation at rather high frequencies, up to about 10^{10} cps. In fact, the relaxation time τ_{rel} for "heating" and "cooling" of carriers is on the order of several times or several dozens of free-path times, meaning that $\tau_{rel} \lesssim 10^{-10}$ sec.

In conclusion, the author expresses his sincere gratitude to V. L. Bonch-Bruyevich, T. M. Lifshits, V.B. Sandomirskiy, and O.I. Yakovlev for valuable discussion of the work.

REFERENCES

1. N.I. Fan', Uspekhi Fiz. Nauk, 1958, 64, 2, 315.
2. J.B. Gunn, Progress in Semiconductors, 1957, 2, 211.
3. H. Brooks, Problemy sovremennoy fiziki, (Problems of Modern Physics) (Translation), 1957, 8, 74.
4. B.I. Davydov, ZhETF (Journal of Experimental and Theoretical Physics), 7, 9 — 10, 1069.
5. Sh.M. Kogan, V.B. Sandomirskiy, Fizika tverdogo tela, (Solid State Physics), 1960, 2, 10, 2,570.
6. E.J. Ryder, Phys. Rev. 1953, 90, 5, 766.

Submitted to the editors 21 December 1960.

---

*The Joule heat lost under these conditions is on the order of a watt.

CERTAIN PROBLEMS OF REFLEX ELECTRON MICROSCOPY

G.B. Der-Shvarts, Yu. M. Kushnir, L.B. Rozenfel'd, et al.

A reflex electron microscope, designed to operate at large observations angles and to obtain diffraction "by reflection" is briefly described. The instrument is provided with an effective illuminating system, an amplifier for the brightness of the final image, a deformation chamber and a device for heating the investigated bulk specimen. The theoretical investigations carried out in connection with the development of the instrument are briefly summarized.

Fert and his co-workers [1] have shown in their time that when the illuminating system in a reflex microscope is inclined by 15 — 25°, it is possible to obtain a fully acceptable image with low-scale distortion and large useful surface of the final image.

We have developed a reflex electron microscope of this type (Figure 1).

It is known [2] that the number of electrons scattered from an object in the range of angles from 15 to 25° relative to the surface of the object, and at very small irradiation angles, is exceedingly small. In this connection, a reflex microscope operating with observations angle 15 — 25° is characterized by low brightness of the final image. It is known [3] that if a reflex microscope is made achromatic with respect to chromatic aberrations and if certain tolerances in the adjustment of the aperture diaphragm are adhered to, the resolution of the microscope is limited only by the chromatic position aberration and by diffraction. The chromatic aberration of position can be reduced by restricting the aperture diaphragm of the objective. This restriction on the dimension of the aperture diaphragm also causes a strong reduction in the brightness of the final image. We were, therefore, faced with the problem of increasing the brightness of the final image for focusing purposes. The final image brightness is increased in our instrument by two ways. First, a more effective electron gun has been developed, and second, a brightness amplifier is used.

Whereas the aperture diaphragm in a transmission microscope plays, for the most part, the role of a contrast diaphragm, in a reflex microscope the function of the aperture diaphragm is indeed in accordance with its name.

Thus, the resolution of a reflex microscope is determined by the dimension of the objective diaphragm, because the radiation reflected from the object surface is quite diffuse. Therefore, the effectiveness of an electron gun in a reflex microscope should be estimated not from the electron brightness (as guns for a transmission microscope are evaluated), but by the density of the current on the object.

Heretofore, all electron guns for microscopes were designed to ensure maximum electron brightness at a minimum load on the object. It is no wonder that we were able to observe that three-electrode guns of special type are capable of producing a much greater current density on the object than the standard guns used in universal microscopes. Unlike the ordinary guns, our design had a conical focusing electrode. We investigated guns with conical electrodes with vertex angles 60, 90, and 120° (Figure 2). The investigation consisted of plotting the current density j at the smallest cross as a function of the anode current I_a. The parameter was the "depth of immersion" h of the end of the cathode filament (the filament diameter was 0.12 mm) relative to the initial conical funnel of the focusing electrode. The bias was adjusted to obtain maximum current under our conditions (U = 60 kv, cathode temperature 2800°K).

The curves of Figure 3 show the optimum results obtained during these experiments.

---

*Reported at the Third All-Union Conference on Electron Microscopy, Leningrad, October, 1960.

For the electrode with angle $\alpha = 120°$ we have h = 0.5 mm; for $\alpha = 90$ and 60°, h = 1.5 mm. For comparison, Figure 3 shows also the curve $j = f(I_a)$ for the ordinary UEM-100 gun, in which the end of the filament was 0.75 mm above the end of the focusing electrode. According to the experiments by Haine and Einstein [4] this position of the end of the cathode is close to optimal for guns of the ordinary type. As can be seen from the curves, the gain in current density amounts to approximately 4.6 times for the gun with $\alpha = 120°$ at a current of 250 microamperes, and 7 times at a current of 500 microamperes. The currents obtained in reflex operation are in the same range.

The gun which we designed was mounted in the illuminating system of the microscope, shown in Figure 4. It has an armored lead and can be mechanically inclined to any angle from 0 to 22°. This inclination is set with a vernier.

Figure 5 shows a section through the objective. A distinguishing feature of the objective is that the pole piece is removed to the object chamber. Thus, various manipulations of the object are possible during the course of observation. The object stage is constructed to permit deformation of metallic objects by tension. A standard dumbbell-shaped specimen was subject to tension with the stress concentrated at the middle. A specimen 3 mm wide and 1.5 — 1 mm thick could be stressed to failure on the stage. The stress concentrator has the form of a cylindrical aperture 1 mm in diameter.

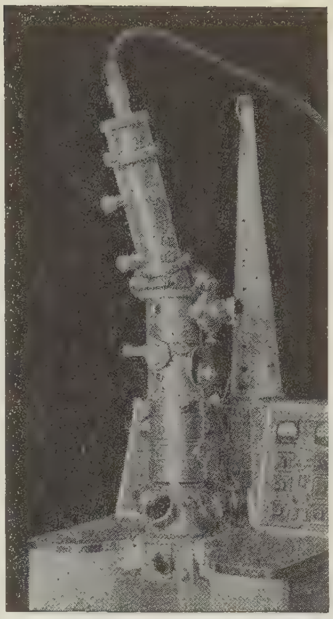


Figure 1. Over-all view of the instrument

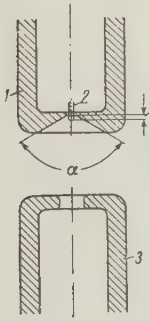


Figure 2

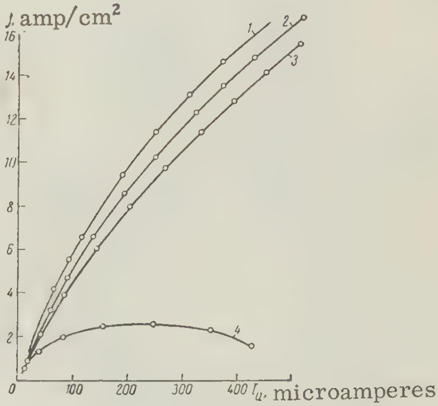


Figure 3

Figure 2. Electron gun with conical electrode: 1 — focusing electrode; 2 — cathode; 3 — anode

Figure 3. Optimum curves showing dependence of current density j in the smallest cross section of the beam on the anode current I_a for conical electrodes with aperture angles α: 1 – h = 0.5 mm, $\alpha = 120°$; 2 — h = 1.5 mm, $\alpha = 90°$; 3 — h = 1.5 mm, $\alpha = 60°$; 4 — h = 0.75 mm

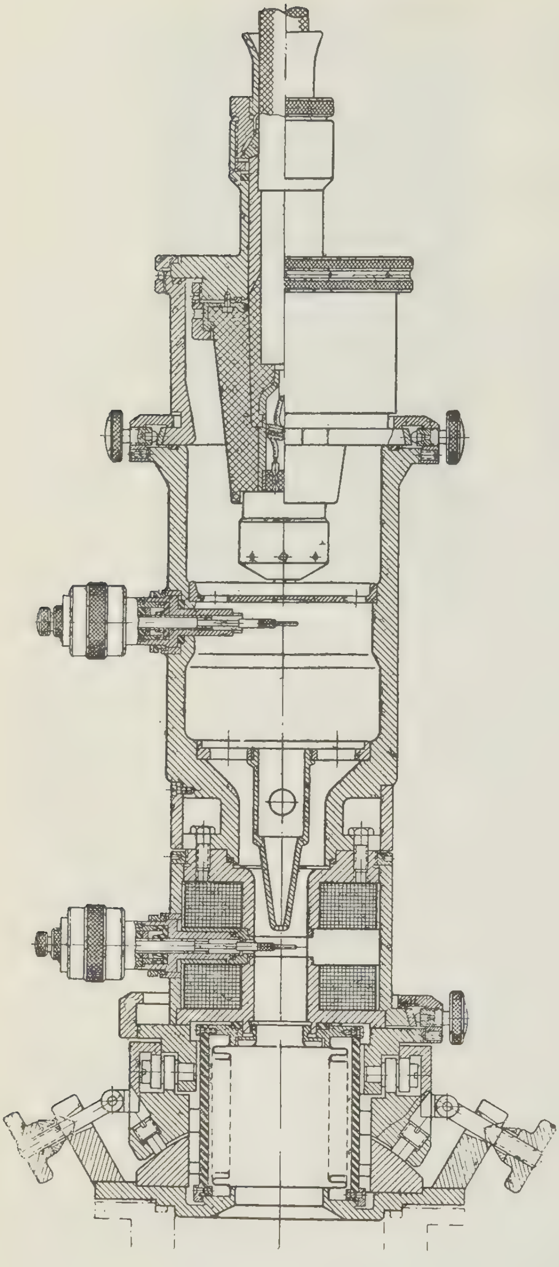


Figure 4. Illuminating system of reflex electron microscope

The same stage also has a holder for an object with standard dimension 10 x 10 mm, and thickness up to 3 mm (Figure 6). Such a specimen can be heated during the observation by means of a special gun, a schematic diagram of which is shown in Figure 7.

The brightness amplifier is mounted in the microscope as shown in Figure 8. Its main purpose is visual observation of the screen during the adjustment of the sharpness of the objective lens. A similar purely electrostatic version of brightness amplifier was described by Zavoiskiy, Butslov, Plakhov, and Smolkin [5], and is used to amplify the brightness of the final image in an electron microscope [6]. The brightness amplifier is mounted under the photocamera of the microscope. The photo magazine of the camera of the microscope is removed from the path of the rays in vacuum. Then the beams forming the image strike an aluminized screen which produces a blue glow. The image produced on this screen is transferred by means of standard specially coated high-aperture photographic optics onto the photocathode of the amplifier. The image on the amplifier screen is viewed through an eyepiece.

The microscope is adapted for diffraction investigations and reflection. During the preliminary investigation of the instrument, we obtained the photographs shown in Figures 9 and 10. Figure 9 shows a photograph from the surface of "silver steel", finished to produce the structure of platelike pearlite. Figure 10 shows a photograph from the surface of white cast iron. The electron-optical magnification is 2500 times, and the resolution is approximately 550 Z.

In parallel with development of the instrument, we investigated theoretically the effect of assembly defects and defects in the shape of pole pieces on the resolution. In view of the fact that the image of the reflex microscope is formed by electrons that have experienced considerable velocity losses, nonalignment of the axis of the pole-piece channels and angular errors in their assembly produce in the near-axis region a constant field perpendicular to the objective axis. This field has an analyzing action and is capable of producing specific chromatic aberrations. Calculation of the aberrations due to the assembly and lens-shape defects calls for a preliminary determination of the corresponding perturbing potentials. These perturbing potentials are usually calculated by the method proposed by Bertein, [7]. For certain problems, this method is mathematically justified from the point of view of analytic continuation. In the form formulated by Bertein, the method is universal in character, and was used by Sturrock [8], Archard [9], and Glaser and Schiske [10] to determine tolerances for the preparation of

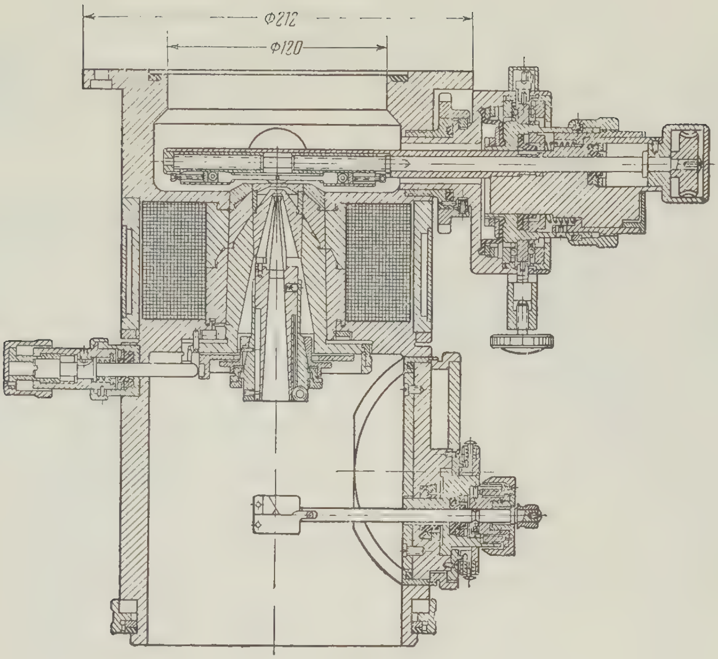


Figure 5. Section through objective of reflex electron microscope

magnetic lenses. It can be shown, however, that this method does not yield exact boundary conditions for the solution of the Laplace equation for the perturbing potentials, for example for the case of a symmetrical elliptic perturbation of the boundaries of the pole-piece channels.

The last problem can be solved only under the assumption that the perturbation is as small as desired. The modified Mathieu functions can be reduced to sums of Bessel functions, individual terms of which are multiplied by the parameter of the Mathieu equation [11]. The fundamental Mathieu functions reduced to sums of trigonometric functions, individual terms of which are also multiplied by this parameter. The latter contains the linear deviation of the semi-axes of the ellipse from circular. We can then calculate the value of the modified Mathieu function on the boundary of the circle inscribed in the ellipse and thus show that the potential on this circle does not coincide with the limiting perturbing potential calculated by the Bertein method.

We have reached the conclusion in our investigations that it is impossible to indicate a general method for the calculation of the perturbing potentials and to determine them in the near-axis region by integrating over a simple layer. By making further assumptions, which reduced the accuracy, we determined the perturbing potentials corresponding to asymmetrical boundary perturbations. Since the scope of this paper does not permit us to discuss the results of that work, we merely note that tolerances for shifts in the pole pieces are on the order of 5 — 10 microamperes for end piece with a geometrical form parameter s/d = 1.5. These pole pieces were used by us in this instrument. The angular inaccuracies in the assembly exert a lesser influence, this being due to the well-known compensation of the perturbing potentials, induced by the opposite ends of the pole pieces. Borsch [12] indicated that chromatic aberrations of all forms can be greatly reduced by using electronic filters.

We have investigated the filter lenses, hoping to increase the resolution of the reflex microscope. The calculations were based on mathematical model of single electrostatic lenses, used by Glaser and Schiske [13] and by Rudenberg [14]. The trajectories in the

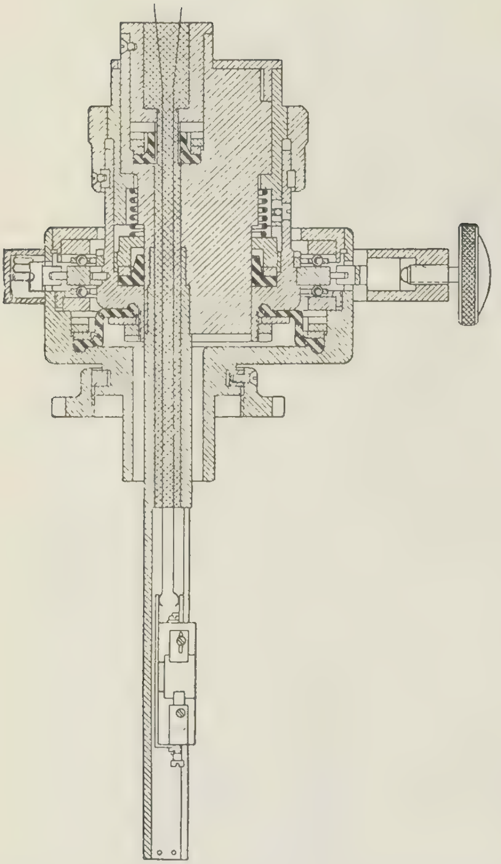


Figure 6. Object stage of reflex electron microscope

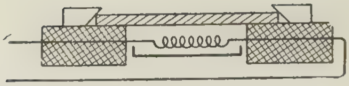


Figure 7. Diagram of electron gun for heating the object

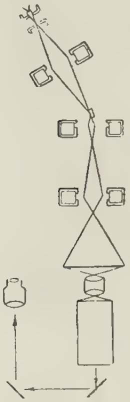


Figure 8. Diagram of brightness amplifier for final image

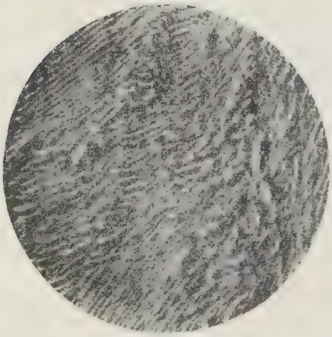


Figure 9. Platelike pearlite, magnification 2500 x.

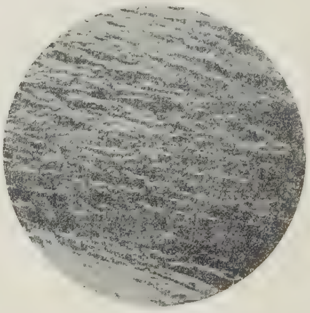


Figure 10. White case iron; magnification 2500 x

fields of these lenses can be expressed analytically, even in the nonparaxial approximation. This facilitates greatly the analysis of the operation of similar lens-filters. The investigations have shown that the resolution of the lenses is limited essentially by the fact that the beams away from the axis, of nonmonochromatic electrons, are focused in different planes. With the electron energy having a spread on the order of 5 — 6 ev, a background is produced on the screen, in which the image produced by the paraxial zones of the field is lost.

REFERENCES

1. C.H. Fert, B. Marty, R. Saporte, Comptes rendus Acad. Sci., 1955, 240, 20, 1975.
2. Yu. M. Kushnir, G.V. Der-Shvarts, Verh. des vierten Internat. Kongress für Elektronenmikroskopie, Berlin, 10—17 Sept. 1958, p. 122.
3. G.V. Der-Shvarts, Yu.M. Kushnir, Radiotekhnika i elektronika, 1959, 4, 6, 1002; G.V. Der-Shvarts, Radiotekhnika i elektronika, 1958, 3, 10, 1315.
4. M.E. Haine, P.A. Einstein, British Journal of Applied Physics, 1952, 3, 2, 40.
5. E.K. Zavoiskiy, M.M. Butslov, A.G. Plakhov, G.E. Smolkin, Atomnaya energiya (Atomic Energy), 1956, 1, 4, 34.
6. Yu.M. Kushnir, V.G. Nyrikov, M.M. Butslov, G.A. Bordovskiy, Pribory i tekhnika eksperimenta (Instruments and Experimental Techniques), 1958, 3, 73.
7. F. Bertein, Ann. radioélectr., 1947, 2, 7, 379.
8. P.A. Sturrock, Philos. Trans. Roy. Soc. London A, 1951, 243, 868, 387.
9. G.D. Archard, J. Scient. Instrum., 1953, 30, 10, 353.
10. W. Glaser, P. Schiske, Z. angew, Phys., 1953, 5, 9, 329.
11. N.W. McLachlan, Theory and Application of Mathieu Functions, Clarendon Press, Oxford, 1947.
12. H. Börsch, Optik, 1949, 5, 7, 436.
13. W. Glaser, P. Schiske, Optik, 1954, 11, 9, 422; 1954, 11, 9, 422; 1954, 11, 10, 455; 1955, 12, 5, 233.
14. R. Rüdenberg, J. Franklin Inst., 1948, 246, 4, 311; 246, 5, 377.

Submitted to the editors 7 February 1961.

THE EM-6 LIGHT-ELECTRON MICROSCOPE

V.N. Vertsner, V.G. Nikiforova, et al.

A combination comprising an electromagnetic single-crystal objective screen of high quality and an optical microscope is incorporated in the simplified original production model of the Em-6 microscope, called the "light-electron" microscope. The electron image produced on the single-crystal screen is viewed through an optical microscope and can be

---

*Reported at the Third All-Union Conference on Electron Microscopy, Leningrad, October, 1960.

photographed by means of an adapter using a "Zenit S" camera. The resolving power of the microscope exceeds 100Å. The accelerating voltage is 35 kv.

The resolving power of the better electron microscope exceeds the resolution of optical instruments by about 200 times. This discloses a possibility of using electron optics at the same degree of magnification to raise the dimensions of the resolvable object to values amenable to observation in optical microscopy.

To bring about the resolution obtained by means of an electron-optical objective, it is necessary to have an intermediate screen of high resolving power. The development of such screens was carried out by V.V. Kuprevich [1]. Crystals obtained by him from fluorite, having dimensions much greater than those of the natural ZnS crystals employed at one time by M. Ardenne [2], have high homogeneity and transparency, while the spectral composition of their glow is close to the sensitivity range of the human eye. The light yield of such crystals differs little from the light yield of polycrystalline phosphors (25 — 30%) of the light yield of polycrystalline screens). The resolution of the screens is approximately 650 — 1000 lines/mm at 35 kv. However, the construction of such a device with a resolution of 10Å is a very complicated problem, since its electron-optical objective must not be inferior in quality to objectives of first-class devices; of equally high quality must be the electric circuitry of the instrument and its mechanisms. Furthermore, crystals having such high resolution (2500 lines/mm) are not yet available. It was, therefore, natural to begin with the construction of a simpler instrument operating on the same principle, proposed by V. N. Vertsner, with a resolution intermediate between that of a light microscope and an electron microscope. At the Second Conference on Electron Microscopy a preliminary communication was made of the possibility of constructing such a microscope [3], which we called a light-electron microscope.

The EM-6 light-electron microscope (Figure 1) is an instrument with an electromagnetic objective, in which the transmitted electron beams produce on a single-crystal high-resolution screen a magnified electronic image which is then viewed through an optical microscope of low magnification and photographed by means of an adapter using the "Zerit S" camera.

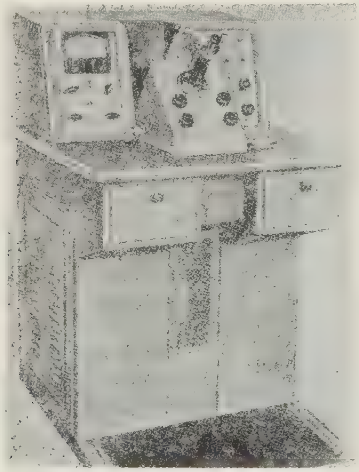


Figure 1. Light-electronic microscope EM-6

The electron source of the EM-6 microscope is a long-focus gun 1 (Figure 2). The cathode filament is fed from the ac line through a transformer located in a high-voltage oil tank. The diaphragm of the anode has a hole 1 mm in diameter. By means of an adjusting device it is possible to center the filament together with the modulating electrode (Wehnelt cylinder) relative to the anode and thereby aim the beam along the optical axis of the instrument.

Directly behind the anode is a centerable condenser diaphragm 2. The mechanism of the diaphragm makes it possible to set successively in position three different diaphragms without breaking the vacuum and to center the diaphragms by moving them in two mutually perpendicular directions.

The illuminating system of the microscope produces, without the use of a condenser lens, a narrow beam of electrons, incident on the object (~ 100 microns).

It should be noted that the use of the light microscope for the subsequent magnification of the image decreases the load on the object by approximately three orders of magnitude compared with an electron microscope of equal final magnification.

The specimens are introduced into the instrument through lock 3. The object, placed in a hollow cylindrical rod, is mounted in the gap between the pole pieces of the objective lens. The rod is sealed on both sides by rubber gaskets and is made up of two parts that are screwed together. When in working position the parts of the rod are screwed apart and the beam of electrons passes freely through the object. To change the object, the rod

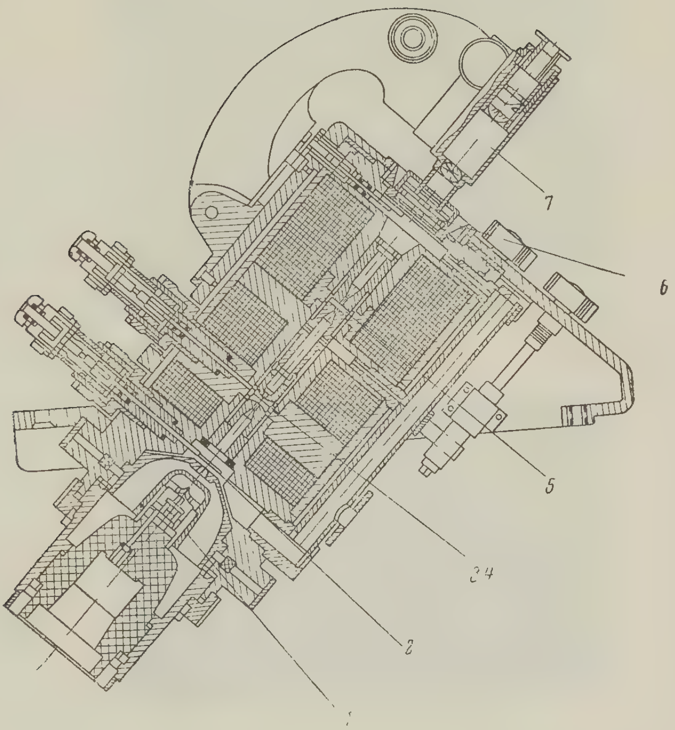


Figure 2. Section through the column of the EM-6 microscope

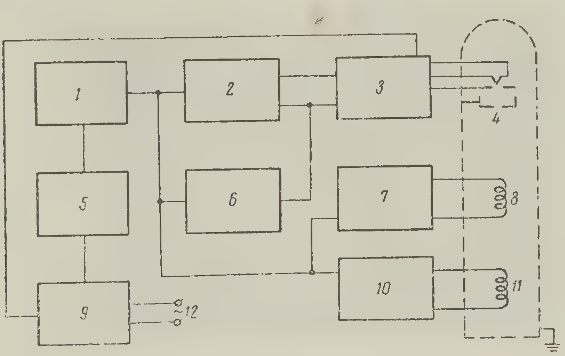


Figure 3. Diagram of electric supply to the EM-6 microscope: 1 — Electronic stabilizer; 2 — generator for high frequency filament circuit of rectifier; 3 — rectifier-multiplier; 4 — electron gun; 5 — rectifier; 6 — high frequency oscillator; 7 — electronic stabilizer; 8 — first objective lens; 9 — ferroresonant stabilizer; 10 — electronic stabilizer; 11 — second objective lens; 12 — line.

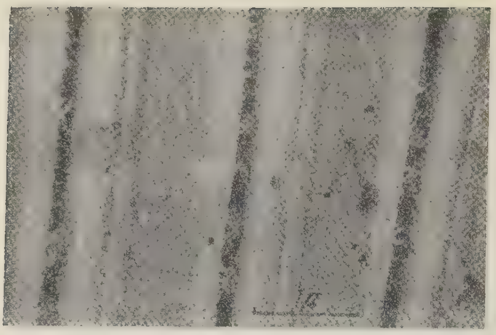


Figure 4. Diffraction grating 600 lines/mm

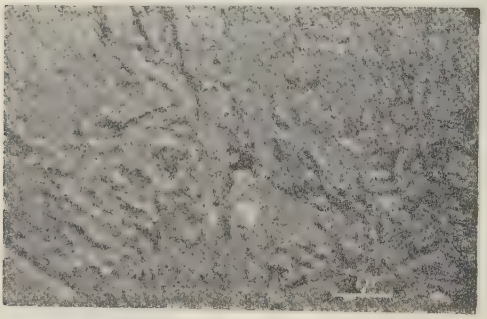


Figure 5. Structure of steel surface; magnification 10,000 ×

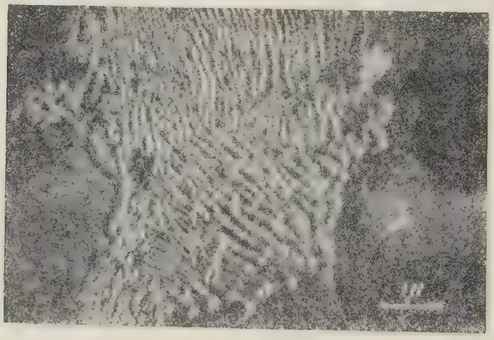


Figure 6. Structure of pearlite surface; magnification 10,000 ×

is screwed together so that the object is in the internal cavity of the rod and is thus protected against contamination as it is pulled through the gaskets into the air, where the rod is again unscrewed and the part of the rod with the object removed from the instrument.

The aperture diaphragm of the objective 4 is also inserted into the gap between the pole pieces of the objective lens, and is placed behind the object. Its mechanism is analogous to that of the condenser diaphragm.

Traveling past the object behind the objective lens, the electron beam enters the second objective lens 5, the magnification of which can be varied in three steps.

The final electron image is formed on a single-crystal screen 6. The side of the screen and of its brass mounting facing the beam are coated with a thin layer of aluminum, so as to prevent charging of the screen. The screen has relatively low thickness (0.4 mm), which is quite adequate in view of the high strength of the fluorite. Since the image obtained on the internal surface of the crystal is viewed through the light microscope, which is located on the outside, the thinner the screen the less distortion it introduces.

The light microscope 7 is mounted on the case of the instrument and can be tilted upward on a hinge, thus making access to the screen possible.

To take photographs, a "Zenit S" camera (without lens) is fastened on the eyepiece of the light microscope; this camera permits both focusing and taking the electron image from the single-crystal screen. The number of pictures per camera load is 36. The most suitable emulsion for photography with the EM-6 instrument is fluorographic film RF-3, but other films with sensitivity 180 — 250 GOST units can also be successfully used, such as type A-2 film.

The exposures range from 2 to 25 sec, depending on the density of the image and on the over-all magnification employed. For an optical-microscope magnification 40×, the over-all magnification is 10,000, 5000, and 2000, respectively. The vacuum system consists of a distributor, VN-494 rotary pump inside the microscope table, and air-cooled NVO-40 diffusion pump. The VKZh-94 silicone oil used in the diffusion pump does not oxidize on contact with air when the oil is hot.

Mounted in the instrument table is a primitive vacuum attachment, which serves to shadowcast the objects.

The instrument is fed from the 220 v AC line through a ferroresonant voltage stabilizer (Figure 3). A high-frequency circuit is used to produce the accelerating voltage. The high frequency oscillator, using a GU-50 tube, generates 8 — 9 kv at 60 kcs. This voltage is fed to a rectifier-multiplier, the output voltage of which reaches 35 kv. The bias voltage on the modulating electrode can be adjusted. The microscope lenses are fed with rectified and stabilized voltage.

The electric circuit of the instrument results in a stable image of good quality.

The EM-6 light-electron microscope is now produced with a resolving power of 150Å on photographs and 80 — 100Å in visual observation. If the pole pieces are carefully manufactured, the resolution can be improved to 60Å on photographs. Further increase in resolution can be attained by stigmation of the objective lens by means of a simple stigmator, since the Fresnel rings are clearly seen in the light microscope. The photographs of Figures 4, 5, and 6 illustrate the quality of the instrument.

It can be hoped that the simplicity of construction and the ease of operation will make this light-electron microscope widely used in research at that important range of magnifications, which is at the borderline between the magnifications of the light-optical and electron microscopes.

REFERENCES

1. V.V. Kuprevich, Inzhenerno-fiz, zh. (Engineering-Physics Journal), 1958, 11, 130.
2. M.F. Ardenne, Elektronenuebermikroskopie, Berlin, I. Springer, 1940.
3. G.A. Bogdankovskiy, V.V. Kuprevich, V.N. Vertsner, I.V. Stepanov, Izv. AN SSSR, Ser. fiz, 1959, 23, 478.

Submitted to the editors 7 February 1961.

ELECTRON-MICROSCOPIC INVESTIGATIONS OF GERMANIUM SURFACE

R.Ya. Berlaga, P.P. Konorov, M.I. Rudenok

Electron microscopic investigations were made of a surface of germanium treated with the standard etchants H_2O_2 and SR-4, with a special etchant that reduces to a minimum the formation of monoxide of germanium and dissolves the germanium dioxide, and with an etchant used to produce homogenous film of monoxide on germanium. A study was also made of the effect of the HF and HNO_3 contained in the SR-4 etchant on the structure of the germanium surface. Parallel measurements were made of the diffusion lengths of the current carriers; the data obtained were compared with the structure of the surface.

The rate of surface recombination and the associated diffusion length of the carriers in germanium are among the basic parameters that determine the operation of various germanium devices (transistors, photodiodes, etc.). These parameters depend strongly on the surface finish of the germanium and vary both as a result of mechanical working of the surface (grinding, polishing, etc.), and as a result of the action of various chemical etchants on the surface. It can be assumed that this is due, either to changes in the surface states brought about by defects of structural character resulting from the mechanical working, or to the appearance of chemically induced films on the germanium surface, principally oxides, of various position, thickness, and degree of homogeneity.

At present there is no clear cut connection between changes in the diffusion length, the rate of surface recombination of carriers, and changes in the state of surface; this limits the possibility of purposefully controlling the surface properties; the choice of etchants to impart specified surface properties is greatly empirical in character.

In the investigation reported here, the surface of germanium was treated with various etchants and investigated with an electron microscope. Simultaneously, measurements were made of the diffusion length of the carriers and an attempt was made to establish a connection between the surface structure of the germanium and the diffusion displacement length.

The research was carried out on single-crystal specimens of germanium of n and p type, with specific resistivity on the order of several ohm-cm, cut along the (111) plane. The electron-microscopic investigations were carried out by the carbon-replica method with an EM-3 microscope operating at 5500 × magnification. The diffusion length was measured by the light-probe method, wherein carriers were injected through a narrow light pipe, and their diffusion displacement length was measured by the drop in the collector signal as the collector was moved away from the light pipe.

We used in our investigation the three most widely used methods for germanium surface finishing: grinding with emery powder, etching in 30% H_2O_2, and etching in the standard etchant SR-4 (50 cc HNO_3, 30 cc CH_3COOH, 30 cc HF, and 0.6 cc Br). Figures 1 and 2 show electron-microscopic photographs of surfaces subjected to grinding and etching in SR-4. The corresponding values of diffusion length were 0.08 — 0.1 and 0.3 mm.

Etching in hydrogen peroxide yielded different surface structures, depending on whether the etching in the peroxide was carried out directly after polishing or after first etching in SR-4. In the former case a well-developed relief is observed, with regular crystal outlines (Figure 3), while in the latter the outward appearance of the surface does not differ from that obtained directly after etching in SR-4. From the corresponding measurements of the diffusion-displacement length it follows that it amounts to 0.22 mm after etching in hydrogen peroxide, independently of the treatment employed prior to this etching (etching in SR-4 or grinding). Thus, an ordered relief does not reduce the diffusion length compared with a smooth surface, whereas a disordered relief, obtained directly after grinding, greatly reduces this length.

---

*Reported at the Third All-Union Conference on Electron Microscopy, Leningrad, October, 1960

Considerable interest is attached to a study of the effect of individual components in compound etchants on the germanium surface. It is assumed that etching in SR-4, which contains HNO_3 and HF, causes the surface of the germanium to be oxidized by the HNO_3 and that the resultant oxide film is dissolved in the HF. There are still no reliable data, however, on the character or structure of this film. We have investigated the action of HNO_3 and HF on the surface of germanium. Before treatment with HF, the germanium was either ground or etched in CP-4. In either case, the surface structure and the diffusion length remained the same after treatment with HF. This confirms that the HF does not interact with the germanium but with its oxide only.

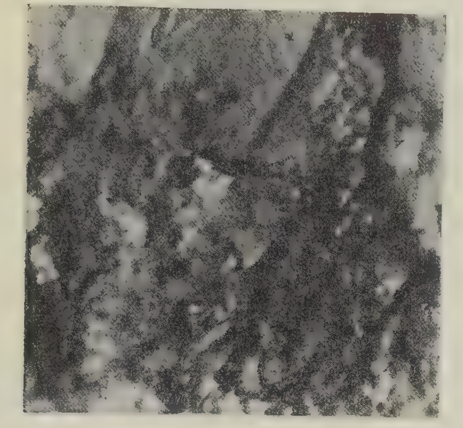


Figure 1. Microphotograph of germanium surface after mechanical grinding

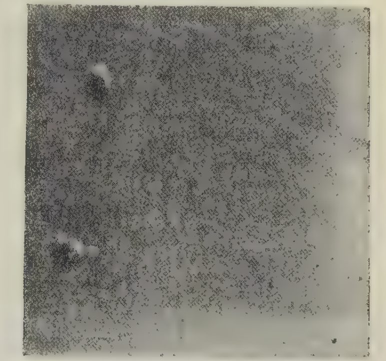


Figure 2. Microphotograph of surface of germanium after etching in CP-4

Figures 4a and 4b show microphotographs of a germanium surface, etched in 63% HNO_3 for 20 minutes after prior etching in SR-4. As can be seen from the microphotographs, the surface of the germanium becomes inhomogenous. Various minute islands are produced on the smooth surface obtained after the prior treatment with SR-4, and these islands consist of grains with a clearly pronounced form of hexagonal pyramids. This gives grounds for assuming that the surface of the germanium becomes covered after treatment with HNO_3 by an inhomogeneous oxide film which, judging from the forming pyramids, can be identified with the hexagonal modification of germanium dioxide. One might thing that the inhomogeneity of the oxide film is due to the fact that it is more intensely produced in places where the surface is damaged. Evidence in this favor is provided by microphotographs of germanium which was first polished and then treated with HNO_3; these photographs show that the germanium surface is covered in this case by a continuous loose film (Figure 4c).

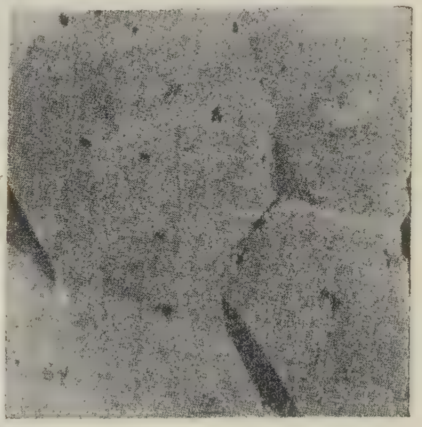


Figure 3. Microphotograph of surface of germanium etched in 30% H_2O_2 after grinding

The measured values of diffusion length of the carriers in germanium first etches with SR-4 and then treated with HNO_3 exhibit a large spread in values, from 0.07 to 0.02 mm, which again confirms the inhomogeneous surface structure. A reduction in the diffusion length compared with values that are typical of a surface treated in SR-4 can be attributed to the formation of an oxide film on the surface, or to inhomogeneities of the surface due to formation of such a film.

In conclusion we give the results of preliminary investigations of the surface of germanium

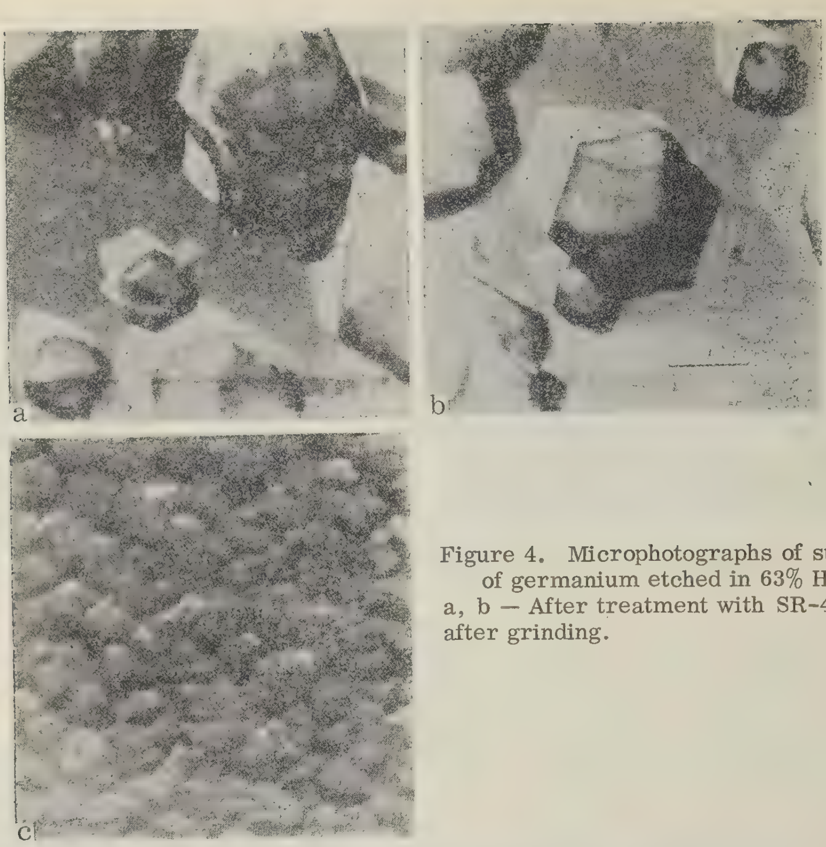


Figure 4. Microphotographs of surfaces of germanium etched in 63% HNO_3: a, b — After treatment with SR-4; c — after grinding.

etched after prior treatment with SR-4 in various etchants, namely etchant No. 8 (20 cc HNO_3 and 10 cc HF), which reduces the production of monoxide to a minimum and dissolves the dioxide, and in etchant No. 5 (40 cc HF, 6 cc H_2O_2 and 24 cc H_2O), used to produce a homogeneous film of monoxide on the germanium. In the former case, the structure and values of the diffusion lengths of the carriers differed little from those corresponding to the initial surface, obtained by etching in SR-4. In the latter case, after etching in etchant No. 5, a more or less homogeneous film is produced on the surface (Figure 5). The diffusion length increases here to 0.5 — 0.7 mm, and consequently the formation of a homogeneous oxide film on the surface of the germanium (at least the monoxide film) can lead to a considerable increase in the diffusion length of the carriers, and it can be assumed that in the case of oxide coatings, the reduction in the diffusion length is essentially due to disturbances in the homogeneity of these coatings.

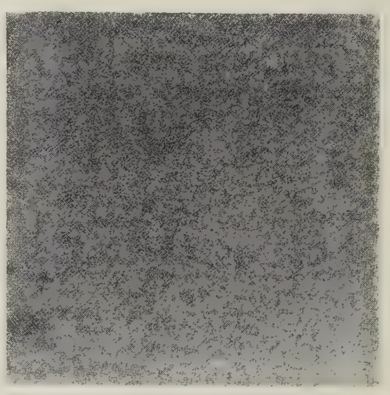


Figure 5. Microphotograph of surface of germanium after etching in etchant No. 5

Submitted to the editors 7 February 1961.

CALCULATION OF THE CHARACTERISTICS OF THE SCATTERING OF MEDIUM-ENERGY ELECTRONS ON FREE ATOMS

A.N. Pilyankevich

Results are reported of a theoretical calculation of the effective cross sections of elastic and inelastic scattering of electrons, which determines the image contrast in an electron microscope. The dependence of the atomic cross section on the atomic number is discussed. The data obtained are compared with the results of Hall's experiment.

INTRODUCTION

To determine the physical properties of an object viewed in an electron microscope by its image contrast it is necessary to establish a quantitative connection between these properties and the contrast, taking account of such structural parameters of the microscope as the accelerating voltage, the illuminator and objective apertures, and so on. Such connections can be established in two ways, either by experiment, calibrating the contrast of the resultant images against previously known object properties, or theoretically. The first path is quite complicated, since a number of factors influencing the contrast is large. It is, therefore, more advantageous to use the theoretical approach, and the results obtained by this method should be rid of inaccuracies due to the assumptions and approximations made by introducing corrections based on experimental data.

Inasmuch as the main process that take place when the beam electrons interact with the substance of the object in the electron microscope is scattering, the purpose of the present work is a theoretical calculation of the effective scattering cross sections σ. We were interested here principally in the cross sections that characterize the probability of scattering of the electrons by an angle close to the effective aperture angles of the objective lens, α_0, since the only electrons participating in the image of the object are those scattered within the aperture angle. The effective aperture angle is determined by the focal distance and by the diameter of the aperture diaphragm of the objective and also by the spherical-aberration constant.

The elastic and inelastic interactions are assumed to be independent processes, so that the cross sections can be calculated for both processes (σ_e and σ_i) independently of each other, assuming the total cross section to be $\sigma = \sigma_e + \sigma_i$.

In the present paper we calculate the cross sections σ_e, σ_i and σ for free atoms, using the Born approximation for electrons of medium energy.

1. ELASTIC SCATTERING

The effective cross section for elastic scattering can be defined as

$$\sigma_e = 2\pi \int_{\alpha_0}^{\pi} \frac{d\sigma_e}{d\Omega} \sin\vartheta \, d\vartheta, \tag{1}$$

where $d\sigma_e/d\Omega$ is the differential elastic-scattering cross section; according to Mott [1], it has the form

$$\frac{d\sigma_e}{d\Omega} = \frac{4}{a_0^2 q^4} [Z - f]^2; \tag{2}$$

where $q = \frac{4\pi}{\lambda} \sin\vartheta$ is the change of the electron wave vector due to the scattering (2ϑ is the scattering angle), and f is the atomic x-ray factor.

---

* Reported at the Third All-Union Conference on Electron Microscopy, Leningrad, October 1960.

Table 1

Cross sections for elastic scattering of 60 kev electrons for atoms with $Z = 3 - 11$ (in units of 10^{-18} cm^2)

Z	Element	$\alpha_0 = 0$	10^{-3}	$2,5 \cdot 10^{-3}$	$5 \cdot 10^{-3}$	10^{-2}	$2,5 \cdot 10^{-2}$	$5 \cdot 10^{-2}$
3	Li	0.450	0.437	0.375	0.252	0.117	0.036	0.008
4	Be	0.715	0.704	0.657	0.508	0.265	0.066	0.020
5	B	0.899	0.889	0.841	0.702	0.411	0.103	0.030
6	C	1.026	1.019	0.939	0.827	0.558	0.152	0.042
6	C_{val}	0.926	—	—	0.804	0.539	0.146	0.035
7	N	1.074	1.068	1.038	0.943	0.685	0.212	0.056
8	O	1.148	1.143	1.116	1.034	0.795	0.277	0.073
9	F	1.190	—	—	1.091	0.870	0.330	0.078
10	Ne	1.224	—	1.204	1.147	0.959	0.420	0.123
11	Na	1.973	—	1.834	1.510	1.086	0.472	—

The elastic scattering cross sections were calculated by numerical integration of (1). We used for the atomic factor f values previously obtained by us [2] and those taken from papers by others [3]. The resultant data are listed in Table 1.

2. INELASTIC SCATTERING

In accordance with Morse's theory [4] the differential cross section of inelastic scattering of electrons has the form

$$\frac{d\sigma_i}{d\Omega} = \frac{4}{a_0^2 q^4} S, \tag{3}$$

where S is the function of incoherent scattering of x-rays. According to the Waller-Hartree theory of Compton scattering [5], the function S is defined as

$$S = Z - \sum_{i,\, j=1}^{Z} |f_{ij}^2|, \tag{4}$$

where

$$f_{ij} = \int \psi_i(\vec{r})\, e^{i\vec{q}\vec{r}} \psi_j(r)\, d\tau. \tag{5}$$

Here $\psi_i(\vec{r})$ are the single-electron wave functions of the scattering atoms. The summation in (4) is over all electrons of like spin.*

In the calculation of the inelastic scattering it is necessary, generally speaking, to use for q the expression

$$q = \frac{2\pi}{\lambda} \sqrt{(2\vartheta)^2 + \left(\frac{\Delta E}{E}\right)^2}, \tag{6}$$

which takes into account the variation in the wave vector not only in direction, but also in magnitude. However, if we assume tentatively the electron energy losses in inelastic scattering to be $\Delta E = 50$ ev and the electron energy to be $E = 100$ kev, the expression (6) becomes equivalent to the formula for the change in the wave vector in elastic scattering even at angles less than 10^{-3} radian.

---

* In earlier investigations of Compton scattering and the inelastic scattering of electrons, all f_{ij} with $i \neq j$ were neglected in the calculation of S. However, recent calculations have shown [6] that an account of these terms brings about a considerable change in the function S.

Table 2

Cross sections for inelastic scattering of 60-kev electrons for atoms with $Z = 3 - 11$ (in units of 10^{-18} cm^2)

Z	Element	$\alpha_0 = 2.5 \cdot 10^{-3}$	$5 \cdot 10^{-3}$	10^{-2}	$2.5 \cdot 10^{-2}$	$5 \cdot 10^{-2}$	Z	Element	$\alpha_0 = 2.5 \cdot 10^{-3}$	$5 \cdot 10^{-3}$	10^{-2}	$2.5 \cdot 10^{-2}$	$5 \cdot 10^{-2}$
3	Li	1.103	0.263	0.084	0.019	—	7	N	0.810	0.401	0.180	0.036	—
4	Be	0.881	0.356	0.115	0.025	0.007	8	O	0.789	0.412	0.195	0.044	0.010
5	B	—	0.380	—	—	—	9	F	—	0.404	0.205	0.049	—
6	C	—	0.390	0.165	0.031	—	10	Ne	0.732	0.389	0.202	0.053	0.012
6	C_{val}	—	0.370	0.155	0.030	—	11	Na	1.248	0.488	0.206	0.058	0.015

Table 2 lists the values of the effective cross sections for the inelastic scattering of electrons with energy 60 kev, obtained by numerical integration of (5) using the same Fok-Hartree self-consistent field functions as were used in the calculation of the atomic x-ray scattering functions and the cross sections for elastic scattering of electrons (the total cross sections for inelastic scattering of electrons were not determined, owing to the lack of sufficiently reliable data for the function S at small scattering angles.

3. TOTAL CROSS SECTIONS. DISCUSSION OF RESULTS

The calculated values of the elastic and inelastic scattering cross sections enable us to determine the total (summary) scattering cross sections, which determine the contrast of the image. Figure 1 shows the total cross sections for different aperture angles as a function of the atomic number of the scattering atom. As can be seen from the plot, the calculated values of the cross sections for each value of the aperture angle can be approximated with sufficient accuracy by the equation

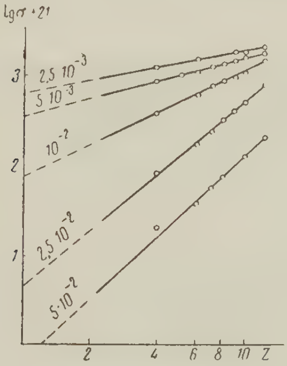


Figure 1

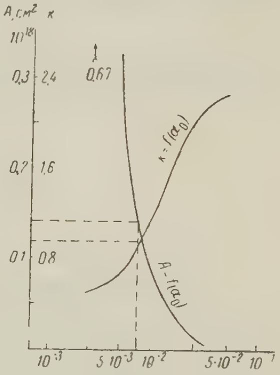


Figure 2

$$\sigma = AZ^k, \tag{7}$$

where A and k are constants that depend on α_0 and E.

The dependence of the constants A and k on the aperture angle for E = 60 kev is shown in Figure 2.

Particular interest attaches to a comparison of our theoretical results with the experimental relation established by Hall [7] for E 65 kev and $\alpha_0 = 5 \times 10^{-3}$ radian. According to this formula

$$\sigma = 0{,}14 \cdot 10^{-18} Z \ (cм^2). \tag{8}$$

It can be seen from Figure 2 that when k = 1 the value of the coefficient A is quite close to 0.14×10^{-18} cm^2, which is precisely the same as in equation (8). The aperture angle here is $5 \times 10^{-3} < \alpha_0 < 10^{-2}$. Thus, the linear dependence of the cross section on the atomic number, established by Hall, is a particular case of a more general power-law dependence (7), and applies only at certain values of the aperture angle of the objective.

CONCLUSIONS

1. The effective cross sections for elastic and inelastic scattering of 60-kev electrons were calculated for nine light elements (Z = 3 — 11).

2. It is shown that the summary cross sections, within the range covered by the investigated elements, can be described by the power-law equation

with constants A and k that depend on the aperture angle and on the accelerating voltage.

3. Good agreement has been obtained between the calculated data and the experimental dependences of the total (summary) cross sections on the atomic number.

REFERENCES

1. N. Mott, Proc. Roy. Soc. A, 1930, 127, 658.
2. A.N. Pilyankevich, ZhTF (Journal of Technical Physics), 1960, 30, 2, 226.
3. P. Morse, Z. Physik, 1932, 33, 11, 443.
4. Internat. Tables for Crystall., 3, 1961.
5. L. Waller, D. Hartree, Proc. Roy. Soc. A, 1929, 124, 119.
6. D.T. Keating, G.H. Vineyard, Acta crystallogr., 1956, 9, 985.
7. C.E. Hall, J. Biophys. and Biochem. Cytol., 1955, 1, 1.

Institute of Metal Ceramics and Special Alloys,
Academy of Sciences, Ukrainian SSR

Submitted to the editors 17 April 1961.

ELIMINATION OF ASTIGMATISM OF THE INTERMEDIATE LENS OF AN ELECTRON MICROSCOPE WITH THE AID OF A STIGMATOR DURING MICRODIFFRACTION

P.A. Stoyanov and N.M. Grishina

We report on the possibility of using a stigmator in the intermediate lens of an electron microscope to improve the quality of the microdiffraction pattern. The investigations were carried out with the UEMB-100 electron microscope. The stigmator comprised two pairs of

---

*Reported at the Third All-Union Conference on Electron Microscopy, Leningrad, October 1960.

coils, located in a plane perpendicular to the optical axis of the microscope. It was placed in the objective tube near the selector diaphragm. The use of the stigmator to eliminate the astigmatism of the intermediate lens resulted in an improvement in the resolution of the electron photographs. No distortion in scale of the diffraction pattern was observed as a result of stigmation.

A comparison of electron photographs obtained in microareas of an object of varying diameters makes it possible to establish thefollowing rule: the smaller the microarea, the higher the resolution of the electron photograph. The improvement in resolution is a result of the small influence of aberration of the intermediate lens, resulting from the decrease in the input aperture of this lens. Both quantities — the dimension of the microarea and the aperture of the intermediate lens — depend on the diameter of the selected diaphragm. Therefore a reduction in the microarea, attained by suitable decrease in the selector diaphragm, is accompanied by a reduction in the aperture of the intermediate lens.

The resolution of electron photographs is predominantly influenced by their spherical aberration and paraxial astigmatism [1]. It is known that the paraxial astigmatism of the intermediate lens can be compensated for with a stigmator [2]. The advantages inherent in using a stigmator are evident when observing electron pictures of large microareas of the object, for in this case the entrance aperture is large and the effect of aberrations considerable. However, the need for observing a diffraction pattern from a large area of the object is a rare event. As a rule, the value of microdiffraction research results increases if the electron pictures cover areas of minimum dimensions. Inasmuch as the entrance aperture of the intermediate lens is small, the effect of aberrations, including astigmatism, is greatly reduced. Whether it is necessary to use a stigmator can be established in this case only by experiment.

One of the tasks of the present investigation was to ascertain the advisability of using a stigmator in the intermediate lens, primarily for the frequently encountered case when the microscope is operated with small intermediate-lens entrance aperture. The investigations were carried out with the UEMB-100 microscope. A stigmator (Figure 1) was located in the

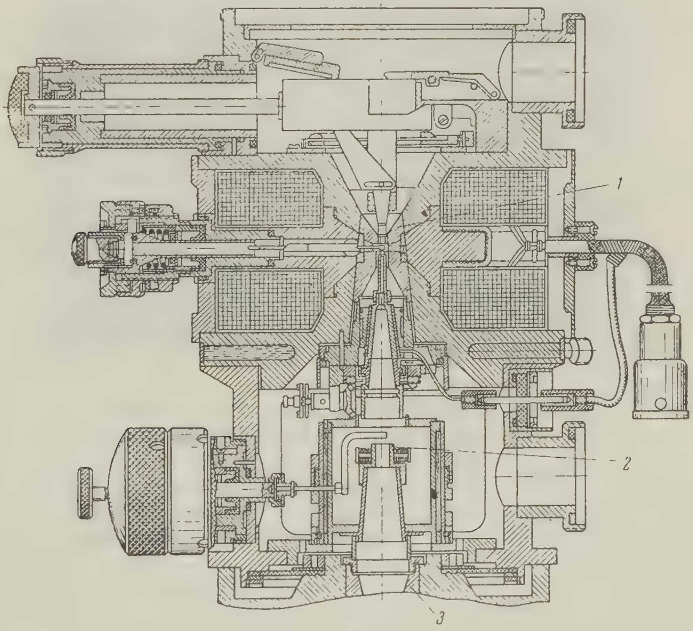


Figure 1. Objective of UEMB-100 microscope with stigmator:
1 — pole pieces of objective; 2 — stigmator of intermediate lens; 3 — pole piece of intermediate lens.

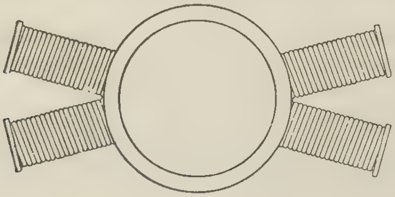


Figure 2. Stigmator of intermediate lens.

tube of the objective, in the direct vicinity of the selector diaphragm. The stigmator consisted of two pairs of coils, located in a plane perpendicular to the optical axis of the microscope. The relative placement of the coils is shown in Figure 2. The windings of each pair of coils were connected in series and were fed independently; a rheostat was used to regulate the excitation current. The geometrical axes of both pairs intersected at a 45° angle. The magnitude and orientation of the astigmatism introduced by the stigmator was varied by suitable adjustment of the current The test object was polycrystalline silver, obtained by thermal sputtering in vacuum on a colloidal film. The electron pictures were observed and photographs taken of object areas of varying sizes. In each case the electron picture was photographed twice, with the stigmator turned on and off. In the former case, the paraxial astigmatism of the intermediate lens was compensated for. The resultant electron photographs of silver are shown in Figure 3. The electron pictures on the left columns were photographed without a stigmator. The upper electron picture was obtained from an area 17 microns in diameter. This is followed by electron pictures from areas 10 and 5 microns in diameter, respectively.

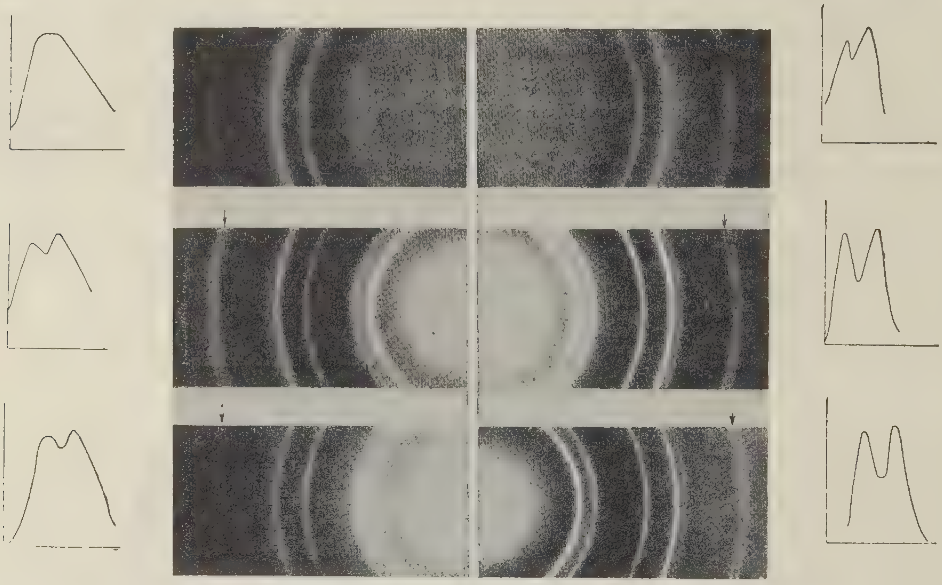


Figure 3. Electron pictures of silver, obtained from different microareas without stigmator (left) and with stigmator (right)

On the right are shown electron pictures photographed from microareas of the same size, but with the stigmator. On any pair of pictures, obtained from one and the same area of the object with the stigmator off and on, one notices an increased resolution of the electron pictures, resulting from the compensation of the astigmatism of the intermediate lens. This is also evidenced by the curves (Figure 3) obtained by photometry of two neighboring rings, marked on the figure by arrows. This pertains to electron pictures obtained from a relatively small area of the object. The smallest diameter of the object portion, for which the electron pictures obtained with and without stigmator were compared, was 2.5 microns. A further considerable reduction in the microportion dimension (below 1 micron) was accompanied by a decrease in brightness of the diffraction pattern, making the observation difficult. Electron pictures taken from an area of minimum dimension are shown in Figure 4. In the case when the input aperture of the intermediate lens was insignificant, amounting to 3 x 10^{-4}, a noticeable deterioration in the resolution of the electron pictures as a result of astigmatism was observed. To be sure, the latter did not always take place. Sometimes the effect of the astigmatism at the indicated entrance aperture was not manifest, since the value of the paraxial astigmatism did not stay constant. It depended not only on the asymmetry of the magnetic field of the intermediate lens, but also on the influence of the electrostatic fields of the charges accumulating on the parts located near the beam. This pertains also to the selector diaphragm — the narrowest place through which the electron beam passes. Obviously, the smaller the diameter of the diaphragm, the stronger the influence of the charges and consequently the greater the astigmatic difference between the focal distances of the intermediate lens. Accumulation of charges has a random character. With time, the formation of contaminations (growth of carbon layers, etc.) causes the magnitude and location of the accumulations to change. This changes the astigmatism of the intermediate lens accordingly.

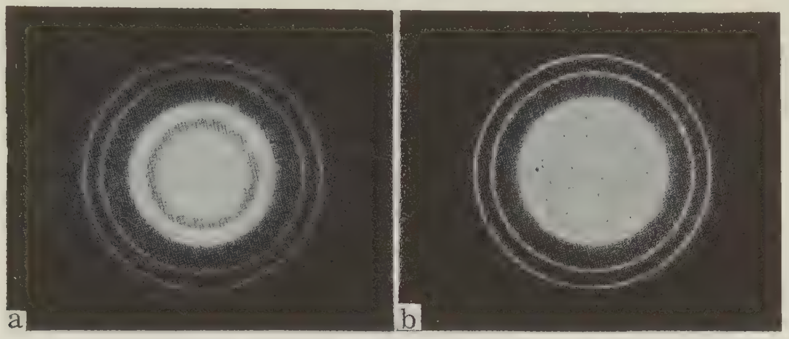


Figure 4. Electron pictures of silver, obtained from a microarea 2.5 microns in diameter
a — without stigmator, b — with stigmator

The caustic method of eliminating astigmatism, which can sometimes be used [2, 3] is not always convenient in our case, since the caustic is observed before the selected diaphragm is placed along the path of the beam. Consequently, in the caustic method of stigmation the influence of the selector diaphragm is not taken into account. If the diaphragm diameter is 1 — 0.4 mm, its effect is small and the results obtained are good. If the diameter is on the order of 0.1 mm and less, the diaphragm can introduce a considerable astigmatism. In this case it is possible to use the caustic for a preliminary adjustment of the stigmator. The final compensation of astigmatism must be carried out after the selector diaphragm is in place. The compensation of astigmatism is established by noting the form of the diaphragm image at the instant when the size of its image passes through a minimum. If the image does not stretch out in any direction up to the minimum itself, the astigmatism is compensated for. Figure 5 shows three photographs of the caustic. The left-hand photograph was taken before compensation of the astigmatism; the central one is taken after the astigmatism was compensated by means of the caustic. The selector diaphragm was then placed in the path of the beam and the final setting of the stigmator effected. After removal of the diaphragm, the caustic had the form shown on the right photograph. Its form indicates the presence of astigmatism. It is obvious that this astigmatism vanishes after the selector diaphragm is returned to its place.

Consequently, the use of a stigmator may be effective also when electron pictures are observed from a small microarea, since the selector diaphragm itself adds to the astigmatism.

The stigmator is a cylindrical lens, which distorts the scale of the image along with the astigmatism, and thus causes the diffraction rings to become elliptic. Our calculations have shown that the distortion in the image scale depends on the placement of the stigmator

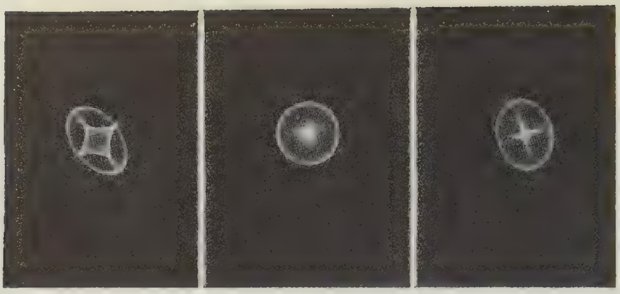


Figure 5. Variation of form of caustic

relative to the selector diaphragm. The numerical value of this quantity can be obtained from the following expression

$$\frac{\Delta r_0}{r_0} = 0.5 \frac{\Delta f}{f^2} t \frac{p+t}{p} .$$

Here $\Delta r_0/r$ is the absolute value of the maximum relative increment in the radius of diffraction rings; Δf is the astigmatic difference of the focal distances, introduced by the stigmator; f is the focal distance of the intermediate lens; t is the distance from the selector diaphragm to the stigmator; p is the distance from the rear focal plane of the objective to the selector diaphragm. Consequently, the closer the stigmator is to the selector diaphragm, the smaller the distortion. Therefore, the selector diaphragm should be located somewhere halfway between the objective and the intermediate lens, where the stigmator should also be placed.

To verify this conclusion, the stigmator was installed successfully in two different places. In the first case it was near the selector diaphragm (Figure 1), and in the second in the pole piece of the intermediate lens. In both cases, the same astigmatic focal-distance differences were set and the diffraction pictures shown in Figure 6 obtained. In order to obtain a sufficiently noticeable distortion of the image scale, the value of Δf in this experiment exceeded the astigmatic focal-distance difference with a compensation of the astigmatism by a factor of six. When the stigmator was located near the selector diaphragm, the distortion of the image scale was minimal (Figure 6a). In the second case (Figure 6b) the distortion of the image scale increased by a factor of several times.

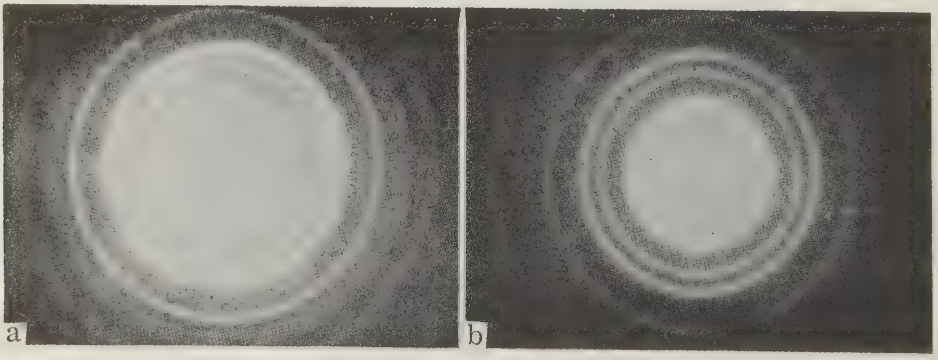


Figure 6. Distortion of scale of diffraction rings when the stigmator is placed near the selector diaphragm (a) and when it is placed in the pole piece of the intermediate lens (b).

Conclusions. The use of a stigmator in the intermediate lens is advantageous not only when electron pictures are taken from large microportions, but also in the practically-important case of investigating diffraction spectra from a portion of objects of minimum size. In either case, the stigmator increases the resolution of the electron pictures.

The stigmator, being a cylindrical lens, distorts the form of the diffraction rings. In order for these distortions to be a minimum, the stigmator must be located in the direct vicinity of the selector diaphragm, and consequently, at a considerable distance from the intermediate lens.

REFERENCES

1. F. Lenz, Optik, 1959, 16, 8, 457.
2. W.D. Riecke, E. Ruska, Z. wiss. Mikrosk. 1957, 63, 5, 288.
3. S. Leisegang, Optik, 1953, 10, 1045.

Submitted to the editors 7 February 1961.

ALIGNMENT OF AN ELECTRON MICROSCOPE BY MEANS OF ELECTROMAGNETIC FIELDS

P.A. Stoyanov

An electron microscope can be aligned by means of deflecting systems that act like prisms.

We consider the properties of a magnetic prism consisting of a cylinder with four cuts, forming two pairs of current leads. Each pair acts as a deflecting system. Results are given of the calculation of the deflection angle as a function of the geometrical and optical parameters, and a formula is derived fro the calculation of the aberrations. The operation of the prism was checked in an electron microscope.

The magnetic field present in the column of an electron microscope result in misalignment of the lens axis and the axes of other optical elements of the instrument. Consequently, high precision in the manufacture of the instrument itself does not in itself ensure suitable alignment of the elements of the column. In microscopes of high resolution, many vital units can as a rule be displaced relative to each other, so that the microscope can be aligned with sufficient accuracy. However, no matter how perfect the alignment mechanisms, such a design reduces the rigidity of the microscope column, and the microscopes become sensitive to vibration. We are faced with the choice of either having a reliable and accurate mechanism to align some unit but reduced rigidity, or increased rigidity at the expense of poor operation of the mechanism.

This alternative has prompted the search for other methods of microscope alignment. In particular, the firm "Metropolitan Vickers" chose a new method for their EM-6 microscopes. The mechanical adjustment of the illuminating system was replaced in this

---

*Reported at the Third All-Union Conference on Electron Microscopy, Leningrad, October 1960.

instrument by magnetic and electrostatic deflecting systems. As a result, the rigidity of the joint between the illuminating system and the objective was retained. Alignment of a microscope with the aid of deflecting field is used also by the "Philips" firm.

Distortions due to aberration occur when deflecting fields, which act like adjustable prisms, are used for alignment. Aberrations of the prisms used to align the illuminating system are not very important. If prisms are used to line the objective, the aberrations may reduce the resolving power of the microscope. The use of prisms in the latter case must, therefore, be separately considered.

In the general case, inaccurate manufacture of the column, or stray fields, cause the axis of the objective to be displaced and turned relative to the axis of the projection lens system (Figure 1). A corresponding tilt of the illuminating system makes it possible to bring either the vertex of the caustic or the magnetic center to the center of the intermediate image [1]. In the former case the microscope operates with minimum spherical aberration, and in the latter with minimum chromatic aberration. In order for both aberrations to be minimized simultaneously it is necessary to make both centers coincident with the center of the image. For this purpose it is possible to place on the path of the beam two adjustable prisms, one near the objective and the other near the object plane of the objective. The effect of the first prism is equivalent to turning the objective in such a way that its axis crosses the axis of the projection system of lenses at a point located in the plane of the lower prism (shown dotted in Figure 1). The effect of the second prism is equivalent to rotating the entire system situated in front of the second prism until the objective axis coincides with the axis of the projection system (shown by a dash-dot line). The result is an aligned system.

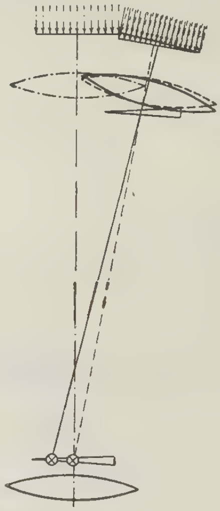


Figure 1. Alignment of objective with the aid of two deflecting systems

The aberrations of the first system should be minimal, for only in this case can we align the microscope without reducing the resolution.

The aberration of the deflecting system can be written in general form as follows [2]:

$$\Delta x_1 = b_8\alpha^2\beta \sin 2\chi - b_4\alpha\beta^2 \cos\chi,$$
$$\Delta y_1 = \frac{1}{2}\,[b_8 + b_{13} + (b_8 - b_{13})\cos 2\chi]\,\alpha^2\beta - b_5\alpha\beta^2 \cdot \sin\chi + b_1\beta^3, \tag{1}$$

where α is half the aperture angle of the beam, β the angle of beam deflection produced by the prism, and b_1, b_4, etc. are aberration coefficients.

To calculate the numerical values of the aberration coefficients, it is necessary to consider specific prism constructions. We give below the results of calculations of a magnetic prism, the construction of which is shown schematically in Figure 2. The prism comprises a cylinder with four cuts, forming two pairs of current leads. Each pair acts as a deflecting system. The magnetic field of the current flowing from point a_1 to a point a_2 deflects the beam in one direction. The field of the current flowing from point b_1 to point b_2 deflects the beam in a perpendicular direction. Independent adjustment of the current in both pairs of current leads with the aid of rheostats R_a and R_b (see diagram in the lower left corner of Figure 2) makes it possible to deflect the beam in arbitrary direction.

A deflecting system of this kind can make use of the manufacturing technology developed for the stigmators of the UEMB-100 and UEMV-100 microscopes [3], and to obtain high accuracy of manufacture and setting of the system in the microscope objective, a necessity for reduction of the aberrations of the system.

The magnetic field of the system, calculated by the Biot and Savart formula in the vicinity of the axis, is

$$H_x = \frac{2\sqrt{2}}{3\pi}\,\frac{I}{c}\,\frac{1}{R}\left(6H_0 - \frac{2H_0 + H_1 - H_2}{R^2}\,x^2 + \frac{2H_0 + H_1 + 5H_2}{R^2}\,y^2 + \ldots\right),$$

$$H_y = \frac{4\sqrt{2}}{3\pi}\,\frac{I}{c}\,\frac{1}{R}\,\frac{2H_0 + H_1 - H_2}{R^2}\,xy + \ldots, \qquad (2)$$

$$H_0 = \frac{l+z}{[R^2 + (l+z)^2]^{1/2}} + \frac{l-z}{[R^2 + (l-z)^2]^{1/2}},$$

$$H_1 = R^2\left\{\frac{l+z}{[R^2 + (l+z)^2]^{3/2}} + \frac{l-z}{[R^2 + (l-z)^2]^{3/2}}\right\},$$

$$H_2 = \frac{3}{2}R^4\left\{\frac{l+z}{[R^2 + (l+z)^2]^{5/2}} + \frac{l-z}{[R^2 + (l-z)^2]^{5/2}}\right\},$$

where H_x and H_y are the x and y components of the intensity of the magnetic field; I is the excitation current; R and L are the radius and half-length of the deflecting system, respectively; c is the velocity of light in vacuum.

The angle of inclination in the prism is $\beta = \theta_n$; θ is a constant calculated from the formula $\theta = \frac{16}{\pi}\frac{1}{c^2}\sqrt{\frac{e}{m}}\frac{I}{\sqrt{u}}$, and n is the ratio of the half-length of the deflecting system to its radius ($n = l/R$). In this, formula, u denotes the accelerating voltage and e/m is the charge to mass ratio of the electron.

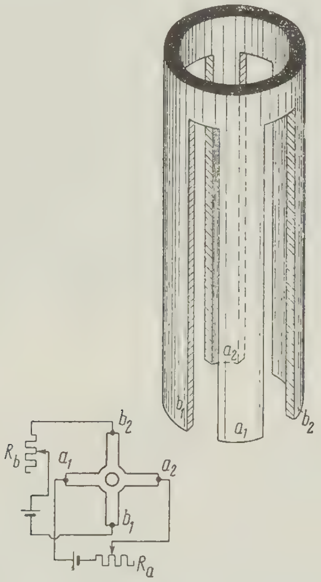


Figure 2

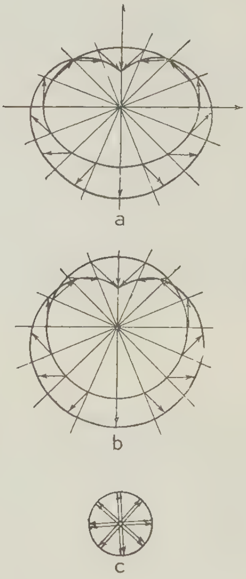


Figure 3

Figure 2. Construction of current leads of the prism. Lower left — electric circuit of independent supply to both pairs of current leads.

Figure 3. Aberration figures:
a — General case; b — transformation of astigmatic ellipse into circle by changing the focus of the objective; c — compensation of astigmatism of the prism by means of the stigmator of the objective. "Degenerate" coma remains (circle).

The aberration in a near-axis point is calculated by the formulas

$$\begin{aligned} \Delta x_0 &= \frac{1}{3} f \left(\frac{f}{R}\right)^2 \alpha^2\beta \sin 2\chi - A(n) f \frac{f}{R} \alpha\beta^2 \cos \chi, \\ \Delta y_0 &= \frac{1}{3} f \left(\frac{f}{R}\right)^2 \alpha^2\beta \cos 2\chi + B(n) f \frac{f}{R} \alpha\beta^2 \sin \chi, \end{aligned} \tag{3}$$

where A(n) and B(n) are coefficients that depend only on n; f is the focal distance of the objective.

The foregoing expressions represent a particular case that follows from the general formulas (1), if we put $b_8 = -b_{13}$ and disregard distortion (the term containing the coefficient b_1), which does not influence the resoltuion of the microscope.

Thus, the deflecting system under consideration introduces an astigmatic aberration of first order (second term), as well as coma, the coma being degenerate (the scattering figure is a circle).

Starting from the results of the foregoing calculation, we establish the following important property of the deflecting system: when the dimensions of the system are proportionally changed (n = const), and the excitation current and accelerating voltage are maintained constant, the angle of deflection β remains unchanged. On the other hand, the aberrations of the system are decreased: the coma is inversely proportional to the square of the radius, and the astigmatism is inversely proportional to the radius itself.

The coefficients A and B can be calculated. If n = 2.5 and R = 3.4 mm (parameters of the investigated system), we get A = 0.75 and B = 0.6. The corresponding aberration figure is shown in Figure 3a. The ellipse is the astigmatic aberration. The arrows show the displacement of the points of intersection between the rays and the image plane, due to the coma. By focusing the objective, it is possible to change the astigmatic aberration figure into a circle (Figure 3b). The stigmator of the objective can compensate the astigmatism, leaving only the coma aberration, also in the form of a circle, shown in Figure 3c.

The operation of the described deflecting system was verified with an electron microscope. The image point of the object, with which the magnetic center coincides, and the magnetic center itself, were displaced together. The vertex of the caustic was displaced by a smaller amount than the point in which the former was located prior to the displacement. This is apparently due to the fact that the position of the zero-coma point depends during the operation of the deflecting system not only on the objective, but also on the deflecting system.

Conclusion. With the aid of a deflecting system, which assume the roles of prisms in electron optics, it is possible to align an electron microscope. The deflecting systems introduce additional aberrations, namely coma and first-order astigmatism, which can be reduced. By using the deflecting system described above, its diameter must be increased, maintaining the ratio of the diameter to the length constant. In this case the angle of declination of the beam, other conditions being equal, remains constant, and the aberrations are reduced.

REFERENCES

1. S. Leisegang, Optik, 1954, 2, 9, 397.
2. W. Glaser, Osnovy elektronnoy optiki, (Fundamentals of Electron Optics) (Russian Translation), GITTL, 1957.
3. P.A. Stoyanov, Optiko-mekhanicheskaya promyshlennost', (Optical-Mechanical Industry) 1958, 4, 40.

Submitted to the editors 7 February 1961.

ANALYSIS OF ABERRATIONS DUE TO DEFECTS IN MANUFACTURE AND IN ALIGNMENT OF ELECTRON MICROSCOPES

Yu.V. Vorob'yev

An analysis of field chromatic magnification and rotation aberrations shows that these can be corrected to an appreciable extent only by precise alignment of the microscopes.

Defects in the manufacture of the objective lens bring about not only axial astigmatism but other aberrations that reduce the resolution of the microscopes. Various methods of microscope alignments are compared, and the requirements that must be satisfied with regard to alignment of microscopes of maximum resolution are discussed.

1. SECOND-ORDER ABERRATIONS

The resolving power of electron microscopes is limited by central aberrations: spherical, chromatic, and diffraction, which in principle cannot be completely eliminated. Along with these aberrations, real instruments are subject to errors due to defects in the manufacture and alignment which prevent attainment of maximum resolution.

Most important of these errors is axial astigmatism of the pole pieces of the objective lens. It is well known that it can be completely eliminated with the aid of a stigmator. However, numerous experiments have shown that when working with pole pieces with a large astigmatic difference of focal distances, the stigmator does not fully compensate for the manufacturing errors and reduces them only by approximately one order of magnitude. This indicates that in a high-resolution instrument with an adjustable stigmator an important role is played by second-order aberrations, resulting from defects in the manufacture of the pole pieces. In the published papers [1, 2, 3] devoted to an investigation of manufacturing faults, the effect of second-order aberration on the resolution was not investigated, since it was correctly assumed that in a system without adjustable stigmator these errors are small compared with the axial astigmatism. After adjustable stigmators were introduced into the high-resolution microscope and the astigmatism eliminated, these errors began to manifest themselves in the quality of the microscope image. An analysis of the most important of these errors, and methods aimed at eliminating them, are discussed here.

Assume that there are no deflecting fields on the axis of the objective and that the astigmatism is compensated for. Then, accurate to second-order terms, the series expansion of the magnetic field has the following form: [4]

$$\begin{aligned} B_z(x, y, z) &= B_z - \frac{1}{4} - B_z''(x^2 + y^2) + \dots, \\ B_x(x, y, z) &= -\frac{1}{2} B_z' x + 6\,\Sigma_3' xy + 3\varepsilon_3'(x^2 - y^2) + \dots, \\ B_y(x, y, z) &= -\frac{1}{2} B_z' y - 6\varepsilon_3' xy + 3\,\Sigma_3'(x^2 - y^2) + \dots, \end{aligned} \tag{1}$$

where ϵ'_3 and Σ'_3 are functions of z and characterize the asymmetry of the field. The corresponding formulas for the aberrations in the Gaussian plane, in the projection of a point on the axis, can be obtained in the usual manner and written in the following form:

---

*Reported at the Third All-Union Conference on Electron Microscopy, Leningrad, October 1960.

$$\Delta x = 3\eta^{*}\alpha^{2}\int\limits_{z_0}^{z_b}[\Sigma_3'\cos(3\theta+2\psi)-\varepsilon_3'\sin(3\theta+2\psi)]\,r_\alpha^3\,dz,$$

$$\Delta y = -3\eta^{*}\alpha^{2}\int\limits_{z_0}^{z_b}[\Sigma_3'\sin(3\theta+2\psi)+\varepsilon_3'\cos(3\theta+2\psi)]\,r_\alpha^3\,dz,$$

where $\eta^* = \sqrt{e/2mv}$; α is the aperture angle, θ the angle of rotation of the trajectory in the magnetic field, and ψ the azimuthal angle in the object plane. Eliminating ψ, we obtain the scattering figure in the form of a circle of radius

$$\Delta r = 3\eta^{*}\alpha^{2}\sqrt{\left(\int\limits_{z_0}^{z_b}(\Sigma_3'\cos 3\theta-\varepsilon_3'\sin 30)\,r_\alpha^3\,dz\right)^2+\left(\int\limits_{z_0}^{z_b}(\Sigma_3'\sin 30+\varepsilon_3'\cos 3\theta)\,r_\alpha^3\,dz\right)^2}. \qquad (2)$$

In order to estimate quantitatively the effect of this aberration on the quality of the image, let us consider the simplest magnetic lens, consisting of a single turn. Let the equations describing the form of the turn have the following form in polar coordinates.

$$\begin{aligned}\varrho &= R_0+\lambda\rho_1(\varphi),\\ \zeta &= \lambda\zeta_1(\varphi),\end{aligned} \qquad (3)$$

where R_0 is the radius of the turn, φ the azimuthal angle, and λ a parameter characterizing the asymmetry. The form of the distorted loop is shown in Figure 1. As can be seen from what follows, the aberration of interest to us is due to distortion of the type

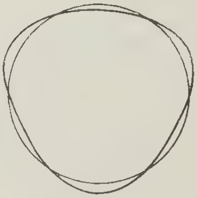


Figure 1. Shape of distorted loop

$$\rho_1(\varphi) = A\cos 3\varphi + B\sin 3\varphi, \quad \zeta_1(\varphi) = a\cos 3\varphi + b\sin 3\varphi. \qquad (4)$$

From the Biot-Savart law we can calculate with the aid of (3) and (4) the perturbation of magnetic field, due to an asymmetry of this kind:

$$\begin{aligned}\varepsilon_3'(z) &= \lambda\frac{5NI\mu_0}{8}\left[-R_0a+\frac{7}{4}\frac{zR_0}{R_0^2+z^2}(za-R_0A)\right]\frac{R_0^2}{(R_0^2+z^2)^{7/2}},\\ \Sigma_3'(z) &= \lambda\frac{5NI\mu_0}{8}\left[-R_0b+\frac{7}{4}\frac{zR_0}{R_0^2+z^2}(zb-R_0B)\right]\frac{R_0^2}{(R_0^2+z^2)^{7/2}}.\end{aligned} \qquad (5)$$

We have numerically evaluated the magnitudes of the aberration by means of formula (5) for the case when the focal distance amounted to 1.2 of the radius of the turn ($j = 1.2R_0$). If we denote by Δl the deviation of the turn from a plane, and by ΔR the radial deviation of the turn from a circle, the values of the aberrations are

$$\Delta r = (3.5\,\Delta l + 1.5\,\Delta R)\,\alpha^2.$$

Taking an illumination aperture of 5×10^{-3}, we find that when the deviation from a circle is 1 micron ($\Delta l = \Delta R = 1\mu$), the aberration is $\Delta r = 1.25$Å.

Thus, the magnitude of this error becomes noticeable in high-resolution microscopes. It can be combatted either by increasing the precision with which the pole pieces are manufactures, or by using a special compensator.

As is done with axial astigmatism, we can construct an auxiliary lens, which compensates for the second-order aberration in the preceding system. If we turn to the formulas for the magnetic field, (1), we readily see that the scalar magnetic potential has the following form in polar coordinates

$$\Phi(x, y, z) = \Phi(z) - \frac{1}{4} B_z' r^2 +$$
$$+ r^3 \left(\Sigma_3' \sin 3\varphi + \varepsilon_3' \cos 3\varphi \right) + \ldots .$$

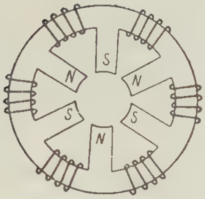


Figure 2. Three magnetic deflecting systems

To cancel out this aberration we must use a magnetic field having similar symmetry properties. Such a field is formed in a system comprising three deflecting systems (Figure 2) in the vicinity of the axis, and has the form

$$\Phi_{\text{K}}(z) = r^3 \Omega'(z) \cos 3(\varphi + \nu),$$

where ν is the angle of rotation of the system relative to the z axis; $\Omega'(z)$ is a function dependent on the configuration of the magnets and proportional to the excitation. It is easy to show that by rotating this system and by choosing the currents in the excitation winding, we can completely correct for the second-order aberration discussed above and thus compensate for defects in the manufacture of the pole pieces of the objective.

2. ALIGNMENT OF ELECTRON MICROSCOPES

The purpose of aligning a microscope is usually to make the axes of all its elements (illumination, objective, intermediate lens and projection lens) coincident. Reference [5] gives calculations and lists tolerances for neutral inclinations of the microscope-lens axes.

Such a formulation of the problem does not appear to be perfectly correct to us. First, we deal here with axes of electron lenses, axes which we cannot see and consequently the attained alignment is very difficult to verify; second, the presence of deflecting fields in the microscope column and in the lenses themselves makes the very concept of lens axes indeterminate.

The practical end purpose of alignment is to correct for all those aberrations that influence the quality of the image. With this, control of the alignment must be carried out against observable quantities, such as the position of the comaless point and the positions of the voltage and magnetic centers. Let us show that these quantities are sufficient to correct for all the nonaxiality aberrations influencing the quality of the image.

As a result of alignment defects, we see on the instrument screen a region of the object which is noticeably remote from the axis of the objective; consequently, the quality of the image is appreciably influenced by nonaxial aberrations. Since the curvature of the image is not appreciable, owing to the smallness of the field of view of the microscope, and since nonaxial astigmatism can, like axial astigmatism, be corrected by a stigmator, two aberrations degrading the image remain, namely coma and field chromatic aberration. We shall dwell on these in greater detail. Let us start with the coma aberration. Let us assume that we are viewing on the screen a certain point P_0 of the object, with corrdinates x_0, y_0. If we denote by x_0', y_0' the tilt of the trajectory in P_0 relative to the objective axis, and by t_x and t_y the projection of the unit vector directed along the illumination axis, and by γ the aperture angle, then

$$x_0' = t_x + \gamma \cos \psi, \quad y_0' = t_y + \gamma \sin \psi.$$

The expressions for the aberrations assume the form

$$\begin{aligned} \Delta x &= B\gamma^3 \cos \psi + \gamma^2 [(2 + \cos 2\psi)(Bt_x + Fx_0 - fy_0) + \sin 2\psi (Bt_y + Fy_0 + fx_0)], \\ \Delta y &= B\gamma^3 \sin \psi + \gamma^2 [(2 - \cos 2\psi)(Bt_y + Fy_0 + fx_0) + \sin 2\psi (Bt_x + Fx_0 - fy_0)], \end{aligned} \tag{6}$$

B, F, and f are the coefficients of the spherical aberrations, the isotropic coma, and the anisotropic coma. We see therefore that there exists on the object a point P_c, in which there is no coma in the case of inclined illumination. We call this the comaless point or sometimes, following [6] the vertex of the caustic. By tilting the axes of the illuminator and of the objective we can shift this point to the center of the screen. Failure to satisfy this condition leads to losses in resolution. Let us estimate the extent of these losses. Let the comaless point be situated a certain distance r from the point P_0, the image of which we see on the screen of the instrument.

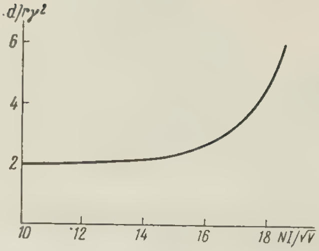


Figure 3. Dependence of coma on the excitation in the objective lens

The condition for the vanishing of the coma in formula (6) yield the following coordinates of the comaless point:

$$x_{\text{к}} = -\frac{B}{F^2+f^2}(Ft_x + ft_y), \quad y_{\text{к}} = -\frac{B}{F^2+f^2}(Ft_y - ft_x).$$

If we put $x_0 = x_c + r_x$ and $y_0 = y_c + r_y$, the point P_0 which we are viewing will be subject in addition to spherical aberration also to the coma error:

$$\Delta x_{\text{к}} = \gamma^2 [(2 + \cos 2\psi)(Fr_x - fr_y) + \sin 2\psi (Fr_y + fr_x)],$$
$$\Delta y_{\text{к}} = \gamma^2 [(2 - \cos 2\psi)(Fr_y + fr_x) + \sin 2\psi (Fr_x - fr_y)],$$

and the magnitude of the coma segment will be

$$d = 3\gamma^2 \sqrt{F^2 + f^2}\, r. \tag{7}$$

The dependence of the resultant coma on the excitation in the objective lens of the microscope is shown in Figure 3. In electron microscopes of high resolution, r must not exceed one or two microns.

Another important error resulting from alignment defects is field chromatic aberration. We have shown earlier [7] that the total field chromatic aberration is the vector sum of the aberrations resulting from the misalignment of the lens axis and the chromatic field aberration proper of the axially symmetrical system. Whereas the latter aberration is negligibly small in microscopes of high resolution, chromatic aberration which results from alignment defects can become quite significant. It is independent of the magnitude or direction of the currents in the projection and intermediate lenses of the microscope, and can be corrected only by alignment. To eliminate field chromatic aberration, it is necessary to shift the magnetic and the voltage centers on the screen by realignment. Unfortunately, owing to the magnetic stray fields and imperfect alignment of the intermediate lens, these centers do not coincide with each other. For accurate coincidence of these centers it is necessary to introduce an adjustable deflecting field between the objective and the projective lens, but at the present time this is not done, it being correctly assumed that adequate magnetic shielding makes the discrepancy between centers minimal. Usually, the alignment of the microscope consists of centering the comaless points and one of the two centers, the voltage center or the magnetic center. The discrepancy between centers must be estimated.

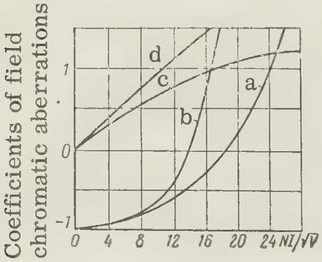


Figure 4. Values of the coefficients of chromatic aberrations of magnification and rotation of the objective (a, d) and also of the chromatic aberrations of magnification and rotation of the projection and intermediate lenses (b, c).

The choice between the centers is dictated by the following considerations. Placing the voltage at the center of the screen eliminates the field chromatic aberration due to oscillations of the current in the objective, and in addition, the image moves along the screen during focusing. Alignment of the magnetic center eliminates these shortcomings, but the drift of the voltage center will then give rise to a field chromatic aberration due to fluctuations in the accelerating voltage. Which of the discrepancies between the voltage and magnetic centers are tolerable and what is the necessary alignment accuracy? We must recall the formulas for field chromatic aberrations [8].

$$\begin{gathered}\Delta r_v = \sqrt{(C_{m_o} + C_{m_i} + C_{m_p})^2 + (C_{\theta_o} + C_{\theta_i} + C_{\theta_p})^2}\, R_v \frac{\Delta V}{V}, \\ \Delta r_i = 2\sqrt{C^2_{m_o} + C^2_{\theta o}}\, R_i \frac{\Delta I}{I},\end{gathered} \qquad (8)$$

where the square root contains the coefficients of chromatic aberrations of magnification and rotation of the objective (o) of the intermediate lens (i), and of the projection lens (p). The values of these coefficients are shown in Figure 4. In formulas (8) R_V and R_i are the distances from the center of the screen to the voltage and magnetic centers (referred to the object), while $\Delta V/V$ and $\Delta I/I$ are the relative oscillations of the accelerating voltage and current.

We note that the coefficient of chromatic aberration of magnification of the intermediate lens, shown in Figure 4, coincides with the corresponding coefficient of the projection aberration only when the intermediate lens operates with magnification.

Having determined the coefficients of chromatic aberrations and knowing the oscillations of the accelerating voltage and current, we can calculate from formulas (8) the chromatic errors and thereby determine the permissible deviations of the magnetic and voltage centers R_i and R_V. Tolerances for the discrepancy between the centers lie within several microns, depending on the required resolution, available stabilization, and magnitude and direction of the currents in the lenses.

REFERENCES

1. W. Glaser, Z. Phys., 1942, 120, 1.
2. P. Sturrock, Philos. Trans. Roy. Soc. London A, 1951, 243, 868, 387.
3. Archard, J. Scient. Instrum., 1953, 30, 352.
4. W. Glaser, Fundamentals of Electron Optics, GITTL, 1957.
5. K. Kanaya, Electrotechn. Laboratory, Tókyo, 1958, 13; J. of Electron Microscopy, 1959, 7.
6. S. Leisegang, Optik, 1954, 11, 49, 397.
7. Yu. Vorob'yev, A. Vyazigin, Doklady AN SSSR, 1959, 126, 6, 1248.
8. V. Vertsner, Yu. Vorob'yev, Optika, Spektroskopiya, 1961, 1, 120.

Submitted to the editors 7 February 1961.

BRIEF COMMUNICATIONS

ON THE CALCULATION OF THE BANDWIDTH IN UHF TROPOSPHERIC SCATTER PROPAGATION

A.V. Prosin

In tropospheric-scatter UHF propagation, the field at the reception point is usually made up of a constant field and of a random field. Under these conditions, the frequency characteristic of the troposphere is a random function of the frequency and of the time. By bandwidth of the troposphere we shall mean that frequency band, within which the inhomogeneity of the random frequency characteristic of the troposphere will not exceed, with a stipulated degree of probability, some prescribed value. In the present communication we continue the work in reference [1] and determine the bandwith of the troposphere in the presence of a constant field and scattered field at the reception point.

Assume that the amplitudes of the signals at frequencies f_1 and f_2 obey the generalized Rayleigh law

$$W_2(u_1,\ u_2) = \frac{u_1 u_2}{\sigma^4(1-R^2)} e^{-\frac{u_1^2+u_2^2+u_{01}^2+u_{02}^2-2u_{01}u_{02}R}{2\sigma^2(1-R^2)}} \times \tag{1}$$

$$\times \sum_{m=0}^{\infty} \varepsilon_m I_m\left[\frac{Ru_1u_2}{\sigma^2(1-R^2)}\right] I_m\left[\frac{u_{01}-Ru_{02}}{\sigma^2(1-R^2)}u_1\right] I_m\left[\frac{u_{02}-Ru_{01}}{\sigma^2(1-R^2)}u_2\right],$$

where u_1 and u_2 as well as u_{01} and u_{02} are the amplitudes of the summary and constant signals at frequencies f_1 and f_2, respectively; σ^2 is the dispersion of the orthogonal components of the random field vector; $R^2 = R_{13}^2 + R_{14}^2$; R_{13} and R_{14} are the correlation coefficients of the orthogonal components of the random field vector; $\epsilon_0 = 1$; $\epsilon_m = 2$ when $m > 0$; I_m is the Bessel function of m-th order of pure imaginary argument. We shall henceforth confine ourselves to the examination of the case $u_{01} = u_{02} = u_0$.

Let us find the probability distribution function for the ratio of the two amplitudes, the distribution of which is described by Equation (1). For this purpose we first determine the probability density $W_2(u_1,\ k_1)$, where $k_1 = u_2/u_1$. The Jacobian of the transformation from the variables u_1, u_2 to the new variables u_1, k_1, will be equal to u_1. Then the two-dimensional distribution function of the quantities u_1 and k_1 will be

$$W_2(u_1,\ k_1) = \frac{k_1 u_1^3}{\sigma^4(1-R^2)} e^{-\frac{u_1^2(1+k_1^2)}{2\sigma^2(1-R^2)}} e^{-\frac{u_0^2}{\sigma^2(1+R_0)}} \times \tag{2}$$

$$\times \sum_{m=0}^{\infty} \varepsilon_m I_m\left[\frac{Ru_1^2k_1}{\sigma^2(1-R^2)}\right] I_m\left[\frac{u_0u_1}{\sigma^2(1+R)}\right] I_m\left[\frac{u_0k_1u_1}{\sigma^2(1+R)}\right].$$

inasmuch as $0 \leqslant u_1 \leqslant \infty$ and $0 \leqslant u_2 \leqslant \infty$, we have $0 \leqslant k_1 \leqslant \infty$.

From formula (2) we obtain for the scattered UHF propagation

$$W_2(u_1,\ k_1) = \frac{k_1u_1^3}{\sigma^4(1-R^2)} e^{-\frac{u_1^2(1+k_1^2)}{2\sigma^2 1-R^2)}} I_0\left[\frac{Ru_1^2k_1}{\sigma^2(1-R^2)}\right]. \tag{3}$$

The expression (3) was derived in reference [1].

The one-dimensional distribution function of the quantity k_1 is

$$W_1(k_1) = \int_0^{\infty} W_2(u_1,\ k_1)\,du_1. \tag{4}$$

To calculate (4) in the case when (2) is used, we expand the Bessel function in the following series:

$$I_m(x) = \sum_{m=0}^{\infty} \frac{(0.5x)^{(m+2n)}}{n!\Gamma(m+n+1)}, \tag{5}$$

where Γ is the gamma function.

When (2) and (5) are used, the calculation of (4) reduces to the determination of the integrals

$$\int_0^\infty t^{\mu-1} I_\nu(\alpha, t)\, e^{-\beta^2 t^2} dt = \frac{\Gamma\left(\frac{\mu+\nu}{2}\right)\left(\frac{\alpha}{2\beta}\right)^\nu}{2\beta^\mu \Gamma(\nu+1)} e^{-\frac{\alpha}{4\beta^2}} {}_1F_1\left(\frac{\nu-\mu}{2}+1,\ \nu+1,\ -\frac{\alpha^2}{4\beta^2}\right), \tag{6}$$

where ${}_1F_1$ is the confluent hypergeometric function, or

$$\int_0^\infty x^n e^{-ax^2} dx = \frac{\Gamma\left(\frac{n+1}{2}\right)}{2a^{\frac{n+1}{2}}}. \tag{7}$$

Formulas (6) and (7) are used when two or three Bessel functions are used, respectively, in the expansion in (2).

If (3) is used, the solution of (4) reduces to a determination of the integral

$$\int_0^\infty e^{-at} I_0(bt)\, t dt = \frac{2a\Gamma(1{,}5)}{(a^2-b^2)^{1{,}5}\sqrt{\pi}}. \tag{8}$$

After calculating (4), we obtain

$$W_1(k_1) = 2(1-R^2)\, e^{-z\frac{2}{1-R}} \sum_{m=0}^\infty \sum_{k=0}^\infty \sum_{l=0}^\infty \frac{\varepsilon_m R^{(m+2k)} z^{(m+l)} \Gamma(2+2m+2k+l)}{k!\, l!\, m!\, \Gamma(m+k+1)\, \Gamma(m+l+1)} \times$$
$$\times e^{z\frac{k_1^2}{1+k_1^2}} \frac{k_1^{(2m+2k+1)}}{(1+k_1^2)^{(2m+2k+2l+2)}} {}_1F_1\left[-(2+m+2k+l),\ (m+1),\ -z\frac{k_1^2}{(1+k_1^2)}\right] \tag{9}$$

when (6) is used and

$$W_1(k_1) = 2(1-R^2)\, e^{-z\frac{2}{1-R}} \sum_{m=0}^\infty \sum_{k=0}^\infty \sum_{l=0}^\infty \sum_{n=0}^\infty \frac{\varepsilon_m z^{(m+l+n)} R^{(m+2k)}}{k!\, l!\, n!} \times$$
$$\times \frac{\Gamma(2+2m+2k+l+n)\, k_1^{(1+2m+2k+2n)}}{\Gamma(m+k+1)\, \Gamma(m+l+1)\, \Gamma(m+n+1)\, (1+k_1^2)^{(2+2m+2k+l+n)}} \tag{10}$$

when (7) is used, where

$$z = \gamma^2 \frac{(1-R)}{(1+R)}; \tag{11}$$

$$\gamma^2 = \frac{u_0^2}{2\sigma^2}, \tag{12}$$

i.e., γ^2 is the ratio of the power of the regular signal to the power of the random signal.

In the absence of a DC component ($u_0 = 0$, $\gamma = 0$, $z = 0$) we obtain from (9) or (10)

$$W_1(k_1) = 2(1-R^2) \sum_{k=0}^\infty \frac{\Gamma[2(k+1)]\, k_1^{(2k+1)}}{\Gamma^2(k+1)\, (1+k_1^2)^{2(k+1)}} R^{2k} \tag{13}$$

or, using (8), we get

$$W_1(k_1) = \frac{2k_1(1-R^2)}{(1+k_1^2)^2 \left[1 - R^2\left(\frac{2k_1}{1+k_1^2}\right)^2\right]^{1{,}5}}. \tag{14}$$

This last formula was derived in [1].

Let us determine the probability of an event wherein the ratios of the amplitudes $k_1 = u_2/u_1 \leqslant K$ or $k_2 = u_1/u_2 \leqslant K$, where $K \leqslant 1$ is a prescribed amplitude ratio. It is obvious that the differential distribution functions of the quantities k_1 and k_2 are equal. Then, taking into account the incompatibility of the simultaneous appearance of ratios k_1 and k_2, we obtain in accordance with the rule for the addition of incompatible events

$$P_1(k < K) = \text{Bep}\{k_1 \leqslant K \text{ or } k_2 \leqslant K\} =$$
$$= \int_0^K W_1(k_1)\, dk_1 + \int_0^K W_1(k_2)\, dk_2 \Rightarrow 2\int_0^K W_1(k)\, dk. \tag{15}$$

Substituting (10) in (15) and integrating, we get

$$P_1(k < K) = 2(1-R^2)\, e^{-z\frac{2}{(1-R)}} \sum_{m=0}^\infty \sum_{k=0}^\infty \sum_{l=0}^\infty \sum_{n=0}^\infty \sum_{s=0}^\infty \varepsilon_m (-1)^s z^{(m+l+n)} \times$$
$$\times R^{(m+2n)} \frac{(m+k+n)!\, \Gamma(2+2m+2k+l+n)\, (1+m+k+l+s)^{-1}}{k!\, l!\, n!\, s!\, \Gamma(m+k+1)\, \Gamma(m+l+1)\, \Gamma(m+n+1)\, (m+k+n-s)!} \times$$
$$\times \left[1 - (1-K^2)^{-(1+m+k+l+s)}\right]. \tag{16}$$

For communication lines with long distances between intermediate stages, most frequently we have $z \ll 1$, corresponding, for example, to $\gamma < 1$ and $R > 0.7$. In this case expression (16) simplifies greatly and can be written

$$P_1(k<K) = e^{-z\frac{2}{(1-R)}}\left[1-\frac{1-K^2}{\sqrt{(1+K^2)^2-4R^2K^2}}\right]. \tag{17}$$

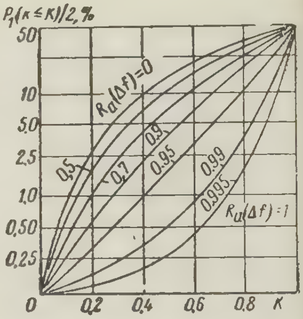
When $\gamma = 0$ we obtain from (17) an expression for the case of scattered UHF propagation, which was derived in [1]. The figure shows the relation $P_1 = (K)$ (17) for $\gamma = 0$ and various $R_a(\Delta f)$, plotted on the basis of the data of [1]. According to [2], when $z < 1$ the connection between R and the frequency correlation coefficient of the signal amplitudes (which can be obtained by experiment) is given by the relation

$$R_{a\gamma}(\Delta f) \simeq R^{2e^{-\gamma^2}} = R_a(\Delta f)^{e^{-\gamma^2}}, \tag{18}$$

where $R_a(\Delta f)$ is the signal-amplitude frequency correlation coefficient in the presence of a random field component at the reception point. On the basis of [3], the quantity

$$R_a(\Delta f) \simeq R^2 = \exp\left\{-\frac{\pi^2}{32}\left(\frac{\Delta f}{\Delta f_2}\right)^2\right\}, \tag{19}$$

where Δf_2 is a certain effective bandwith of the troposphere when directional antennas are used. The full expression for Δf_2 and its dependence on various parameters are discussed in reference [4].



Let us consider as an example the determination of the bandwith of the troposphere for $\gamma = 0$. It is stipulated that the inhomogeneity of the frequency characteristic not exceed 3 db within the band ($K = 0.7$) during 85% of the time. From the figure and from formula (19) we get

$$\Delta f \simeq 0.4\,\Delta f_2. \tag{20}$$

When $d = 300$ km, $R_e = 8500$ km, $p = 1$, $n_1 = 2$, $a = 1$, and $c = 3 \times 10^5$ km/sec (the symbols are all defined in [4]), the bandwidth of the troposphere is $\Delta f \simeq 550$ kc.

From an analysis of (17), (20), and the figure it follows that the bandwith of the troposphere increases with increasing effective band, intensity of the constant wave, and coefficient of the frequency correlation.

REFERENCES

1. I. A. Gusyatinskiy, Bandwidth and Power of Transient Noise in Tropospheric Scatter Radio Communication, Elektrosvyaz' (Electrical Communication) 1959, 4, 3.
2. I. A. Gusyatinskiy, Signal Distortion in Beyond the Horizon UHF Propagation, Ib, Trudov, NII, MS, SSSR (Collected Papers of the Scientific Research Institute of the Ministry of Communication USSR), 1959, 1 (15), 15).
3. A. V. Prosin, V. G. Gubskiy, On the Theory of Separated Reception in Long-Distance Tropospheric Propagation of UHF, Radiotekhnika, 1959, 4, 5, 23.
4. A. V. Prosin, On the Calculation of Cross Noise Power in Long-Distance Communication Systems, Radiotekhnika i elektronika, 1961, 6, 1, 14.

Submitted to editor March 27, 1961

ON THE SYNTHESIS OF AN ANTENNA WITH CIRCULAR APERTURE

B.M. Minkovich, Yu.I. Davidchevskiy

In the realization of directivity patterns of the "pencil" type, the principal role is played by antennas with circular apertures. So far, however, methods for the synthesis of the amplitude distribution for a specified directivity pattern of a circular-aperture antenna have not been sufficiently well developed. We describe below a new method for the synthesis of such a distribution.

From the point of view of the directivity pattern in one of the planes, a flat aperture can be replaced in the aperture method, by an equivalent linear antenna, the amplitude distribution of which is determined by the shape and amplitude distribution of the antenna with the circular aperture [1]. For a circular aperture with symmetrical inphase amplitude distribution, the equivalent antenna is the same for all planes, and the amplitude distribution of this antenna can be determined from the following expression:

$$I_1(\xi) = \int_0^{\sqrt{1-\xi^2}} A(\rho)\, d\eta, \tag{1}$$

where $I_1(\xi)$ is the amplitude distribution of the equivalent linear $A(\rho)$ is the amplitude distribution of an antenna with circular aperture; ξ, η — the reduced Cartesian coordinates in the aperture; $\rho = \sqrt{\xi^2+\eta^2}$; $\eta = \sqrt{1-\xi^2}$ is the equation of the envelope of the aperture (circle).

At the present time the procedure for the synthesis of the amplitude distribution of linear antennas for a specified directivity pattern has been sufficiently fully developed in many papers, for example in [2]. Therefore, in the case when $I_1(\xi)$ can be determined by starting from a specified three-dimensional directivity pattern of an antenna with round aperture, Equation (1) can be used to determine $A(\rho)$.

We shall henceforth assume $I_1(\xi)$ known.

1. After making the change of variable $\xi^2 + \eta^2 = t$, Equation (1) becomes

$$\frac{1}{2}\int_{\xi^2}^{1} A(\sqrt{t})\, \frac{dt}{\sqrt{t-\xi^2}} = I_1(\xi). \tag{2}$$

Equation (2) is a particular case of the integral equations of the first kind, the Abel equation. As is well known [3], the Abel equation can be solved in general form. Using the ordinary procedure, we obtain the solution of equation (1) in the following form:

$$A(\rho) = \frac{2}{\varphi\Pi}\left[\frac{I_1(1)}{\sqrt{1-\rho^2}} - \int_{\rho}^{1} \frac{I_1'(y)\, dy}{\sqrt{y^2-\rho^2}}\right], \tag{3}$$

where ρ enters into the expression under the integral sign as a parameter.

We note that formula (3) is valid for functions $I_1(\xi)$ which are symmetrical with respect to the axis $\xi = 0$, even if the variable ξ is contained in $I_1(\xi)$ to an odd power. For example, for $I_1(\xi) = 1 - |\xi|$ the corresponding distribution is

$$A(\rho) = \frac{2}{\varphi\Pi} \ln\left(\frac{1}{\rho} + \sqrt{\frac{1}{\rho^2} - 1}\right).$$

The integral in turn in (3) drops out if $I_1(\xi) = \text{const}$. Consequently, a distribution of the type

$$\frac{I_1(1)}{\sqrt{1-\rho^2}} \tag{4}$$

produces in a circular aperture an effective uniform distribution, i.e., the directivity pattern of a round-aperture antenna with amplitude distribution given by (4) coincides with the directivity pattern of a linear antenna with uniform distribution.

The term outside the integral sign in (3) differs from zero only when $I_1(\xi)$ has at $\xi = 1$ a "pedestal," i.e., $I_1(1) \neq 0$. Thus, expression (3) makes it possible to calculate easily the amplitude distributions of antennas with round aperture, analogous to the distributions of linear antennas.

2. Let us calculate by the method described above the quasi-optimal distribution of an

antenna with round aperture. The expression for the quasi-optimal distribution of the equivalent linear antenna will be taken in the form proposed in reference [4]. Then, using in (3) the representation of $I_1(\xi)$ from reference [4], we obtain an expression for the quasi-optimal distribution of an antenna with round aperture

$$\frac{A(\rho)}{A(0)} = \frac{\beta^3}{\beta \operatorname{sh} \beta - \operatorname{ch} \beta + 1} A_1(\gamma), \tag{5}$$

where β determines the level of the side lobes ($\cosh \beta$ is the relative level of the side lobes of the optimal antenna);

$$A_1(\gamma) = \frac{\gamma \operatorname{sh} \gamma - \operatorname{ch} \gamma + 1}{\gamma^3}; \tag{6}$$

$$\gamma = \beta \sqrt{1 - \rho^2}. \tag{7}$$

A plot of the function $A_1(\gamma)$ is shown in Fig. 1.

The function $A_1(\gamma)$ is actually a nonnormalized quasi-optimal amplitude distribution, but the abscissas represent not ρ, but a quantity γ, determined by β and ρ. Part of the plot located to the left of the value $\gamma = \beta$ represents the amplitude distribution of an antenna with circular aperture, which tends to ∞ as ρ approaches unity ($\gamma \to 0$).

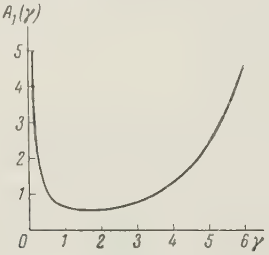


Fig. 1

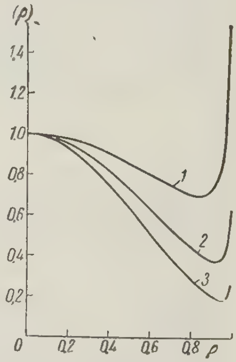


Fig. 2

In Fig. 2, curves 1, 2, and 3 are the normalized values of the amplitude distribution as functions of ρ for three values of β, namely 3, 4.2 and 5.3, respectively. These values yield for an optimal antenna a side-lobe level of 20, 30.4 and 40 db or a level 6 db smaller for the quasi-optimal antenna. The curves of Fig. 2 extend to $\rho = 0.99$; in the interval from 0.99 to 1, $A(\rho)$ increases to infinity. A calculation of the curves after plotting $A_1(\gamma)$ reduces to breaking up the scale of γ for different values of β. These curves are similar to those from reference [5], with the exception, naturally, of the values on the edge. Although the quasi-optimal distribution of a circular-aperture antenna has an infinite value on the edge, and this cannot be realized in practice, the influence of this on the directivity pattern can be estimated.

The method described above enables us to interrelate linear antennas and antennas with circular apertures; distributions "characteristic" of linear antennas can be applied to those with circular apertures, and conversely, expression (1) makes it possible to apply to linear antennas the amplitude distributions corresponding to the directivity patterns that are "characteristic" of antennas with circular aperture.

We note that formula (3) is valid also for out of phase symmetrical amplitude-phase distributions.

REFERENCES

1. G. F. Koch, Flachenstrahler mit kleinen Nebenmaxima, Fernmeldetechn. Z., 1954, 7, 10, 498.
2. G. H. Brown, Pattern synthesis — simplified methods of array design to obtain a desired directive pattern, RCA Rev., 1959, 20, 3, 398.
3. I. I. Privalov, Integral Equations, ONTI, 1935.
4. B. M. Minkovich, On the question of quasi-optimal linear inphase antennas with continuous current distribution, Radiotekhnika i elektronika, 1959, 4, 6, 1057.

5. R. C. Hansen, Tables of Taylor distributions for circular aperture antennas, IRE Trans., 1960, AP-8, 1, 23.

Received by editor June 30, 1960

TRANSIENTS IN A TWO-CIRCUIT PARAMETRIC AMPLIFIER

Yu. A. Il'inskiy, and M. D. Karasev

1. Analysis of Transients

The stationary modes of a two-circuit parametric amplifier was considered by Manly and Rowe, by Heffner and Wade, and by others [1 — 5]. In these papers, use was made of the energy relations derived by Manly and Rowe [1], and the principle of harmonic balance under small-signal conditions was used for a more detailed analysis.

In the present paper we consider transients in a parametric amplifier under the assumption of a small signal, i. e., the settling of the oscillations and the operation of an amplifier in response to a pulse signal.

Certain two-circuit parametric amplifiers are shown in Fig. 1. These are systems with two degrees of freedom and one periodically varying reactance element.

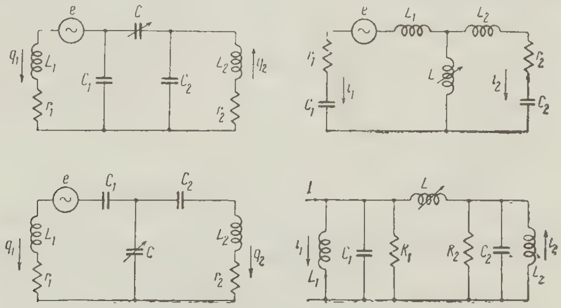


Fig. 1. Diagrams of two-circuit parametric amplifiers.

Using suitable coordinates (see Fig. 1) we can write the equations for such systems in the form

$$\ddot{x}_i + \omega_i^2 x_i = \mu\left(-2\delta_i \dot{x}_i + \alpha_i f(\omega t)(x_1 + x_2) + e_i\right). \tag{1}$$

Here i = 1, 2 is an index pertaining to the first or second loop; e_1 is the external voltage; $e_2 = 0$; μ is a small parameter; δ is the damping in loop i; f is a periodic function with period 2π; ω is the frequency of variation of the parameter.

As follows from the form of Equation (1), the damping of the resonant circuits and the depth of modulation of the reactive coupling are assumed to be small. Thus, the Q of the circuits is assumed large and the parametric coupling between circuits is assumed weak. Then the oscillations in the system will be close to harmonic. As regards the constant coupling between the circuit, it may also be large; in this case it is sufficient to change over to the normal coordinates of a system that differs from the initial system in that the active resistances and the parametric couplings are eliminated. In this case one obtains equations that are somewhat more complicated than (1), but the abbreviated equations of the new system are of the same form as the abbreviated equations of system (1).

Equations (1) can be solved by one of the quasi-linear methods. The most convenient is the averaging method [6]. To employ this method it is first necessary to reduce Equation (1) to standard form.

Let us consider first the case of the so-called inverting modulator [5], i. e., let $\omega \simeq \omega_1 + \omega_2$. To reduce the equations to standard form, we first assume, as is customary,

$$x_i = Z'_i e^{j\omega_i t} + Z'_{-i} e^{-j\omega_i t},$$

$$\dot{x}_i = j\omega_i \left(Z'_i e^{j\omega_i t} - Z'_{-i} e^{-j\omega_i t} \right),$$

$i = 1$ or 2; $Z'_{-i} = Z'^{*}_{i}$ are complex amplitudes.

Let $e_i = E_i e^{jpt} + E_{-i} e^{-jpt}$, where $E_1 = E$; $E_2 = 0$; p is the frequency of the external force ($p \simeq \omega_1$); E is the slowly varying complex amplitude of the external voltage. We put $\mu\Delta_1 = p - \omega_1$, $\mu\Delta_2 = (\omega - p) - \omega_2$, and introduce a change in the complex amplitudes $Z'_i = Z_i e^{j\Delta_i t}$. We then obtain the standard equations for Z_1. Averaging of the right halves of the resultant equations yields the symbolic equations

$$\dot{Z}_1 = \mu \left\{ -\left(\delta_1 + j\Delta_1 + \frac{j\alpha_1 f_0}{2\omega_1} \right) Z_1 - \frac{j\alpha_1 f_1}{2\omega_1} Z_{-2} + \frac{E}{2j\omega_1} \right\},$$
$$\dot{Z}_{-2} = \mu \left\{ -\left(\delta_2 - j\Delta_2 - \frac{j\alpha_2 f_0}{2\omega_2} \right) Z_{-2} + \frac{j\alpha_2 f_{-1}}{2\omega_2} Z_1 \right\}, \tag{2}$$

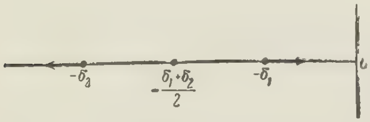


Fig. 2. Characteristic roots in the case of optimal tuning.

where f_0, f_1, and f_{-1} are the Fourier coefficients of f (ωt);

$$f(\omega t) = \sum_{k=-\infty}^{k=+\infty} f_k e^{j\omega k t}.$$

We shall assume that $f_0 = 0$, so that we can introduce $\Delta'_i = \Delta_i + \frac{\alpha_i f_0}{2\omega_i}$ and thus take into account the detuning due to the constant component of the coupling. The solution of the linear equations (2) with constant coefficients can be readily found for any external voltage E and for all initial conditions. Thus, we can find the transient mode, i. e., the establishment of the oscillations in the system when E = const and under the zero conditions $Z_i(0)$ [i.e., $x_i(0) = x_i(0)$]. The equations are in matrix form

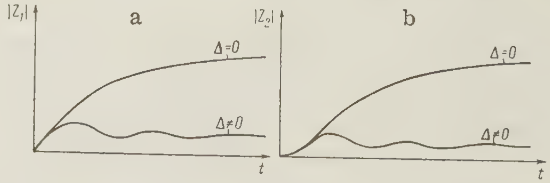


Fig. 3. Settling of the amplitude in the first (a) and second (b) circuits.

$$\dot{Z} = AZ + B,$$

$$\dot{Z} = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}, \qquad B = \mu \frac{E}{2j\omega_1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The solution, satisfying the initial conditions Z(0) = 0, is [7]

$$Z = \int_{1}^{t} e^{A(t-s)} B(s)\, ds = \tag{3}$$

where $\mu\gamma_{1,2}$ are the characteristic numbers of the matrix

$$\mu \begin{pmatrix} -(\delta_1 + j\Delta_1) & -\frac{j\alpha_1 f_1}{2\omega_1} \\ \frac{j\alpha_2 f_{-1}}{2\omega_2} & -(\delta_2 - j\Delta_2) \end{pmatrix}.$$

When t → ∞, we obtain a stationary solution that coincides, for example, with those published in [3].

Solution (3) can be written in more detailed form

$$Z_1 = \frac{E}{2j\omega_1(\lambda_2 - \lambda_1)}\left[-\frac{\delta_1 + j\Delta_1 + \lambda_2}{\lambda_1}(1 - e^{\mu\lambda_1 t}) + \frac{\delta_1 + j\Delta_1 + \lambda_1}{\lambda_2}(1 - e^{\mu\lambda_2 t})\right],$$

$$Z_{-2} = \frac{E}{2j\omega_1(\lambda_2 - \lambda_1)}\,\frac{j\alpha_2 f_{-1}}{2\omega_2}\left[\frac{1 - e^{\mu\lambda_1 t}}{\lambda_1} - \frac{1 - e^{\mu\lambda_2 t}}{\lambda_2}\right]. \tag{4}$$

We recognize that

$$x_i = Z_i e^{j\omega_i t}\, e^{j\mu\Delta_i t} + \text{complex conj.} = 2\,|Z_i|\cos[(\omega_i + \mu\Delta_i)t + \arg Z_i].$$

Here $|Z_i|$ and $\arg Z_i$ are the slowly varying amplitude and phase of the oscillations with frequency $\omega_i + \mu\Delta_i$, i.e., p and $\omega - p$.

The following can be concluded from equations (3) and (4):

1) In a two-circuit parametric amplifier, the settling time does not depend on the relation between the phase of the suddenly applied signal and the phase of the pulsations of the parameter. The phase of the oscillations in the first circuit is also independent of the phase of the parameter pulsations, and the phase of the oscillations in the second circuit is made up of the phases of the input signal and the parameter.

2) Under optimum tuning ($\Delta_1 = \Delta_2 = 0$) the characteristic roots $\lambda_{1,2}$ are real and equal to

$$\lambda_{1,2} = -\frac{1}{2}\left[(\delta_1 + \delta_2) \pm \sqrt{(\delta_1 - \delta_2)^2 + \frac{\alpha_1\alpha_2}{\omega_1\omega_2}|f_1|^2}\right].$$

When $f_1 = 0$ (there is no change in parameter) they are naturally equal to $-\delta_1$ and $-\delta_2$, and as $|f_1|$ increases, they diverge symmetrically with respect to the center of the segment $(-\delta_2, \delta_1)$ (see Fig. 2). When the root with the smallest absolute value reaches zero, the system becomes self-excited. When this root decreases in absolute value, the duration of the transient increases ($\sim 1/-\lambda$). As follows from (4), the transient process in the case when $\Delta_1 = \Delta_2 = 0$ is the sum of two exponentials in the first circuit and the difference of two exponentials in the second circuit; with this $\left.\frac{dz_{-2}}{dt}\right|_{t=0} = 0$. The transient has accordingly the form shown in Fig. 3a for the first circuit and in Fig. 3b for the second circuit. As regards the duration of the transient, it is determined, in the case of large regeneration, by the characteristic number having the smallest absolute value, and is proportional to the voltage gain.

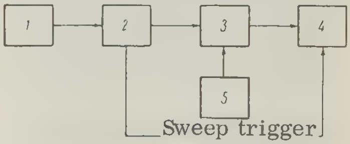


Fig. 4. Block diagram of experimental setup:

1 — Signal generator; 2 — modulator; 3 — parametric amplifier; 4 — oscillograph; 5 — pumping generator.

3) If one of the circuits is detuned (Δ_1 or $\Delta_2 \neq 0$), then $\operatorname{Im}\lambda_{1,2} \neq 0$ and the transient contains damped oscillations (see Fig. 3). If with this $\Delta_1 + \Delta_2 = 0$, i.e., only the external signal is detuned, then the frequency of these oscillations is equal to the detuning Δ_1 of the external signal. On the other hand, if $\Delta_1 = 0$ but $\Delta_2 \neq 0$, the pumping frequency is detuned, then the oscillation frequency in the transient process is less than the detuning, and is much less if $\delta_2 \gg \delta_1$, as usually takes place.

Analogous formulas can be obtained also for a noninverting modulator from the equations for $Z = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$.

2. EXPERIMENT

We observed experimentally the transient processes in a balanced parametric-amplifier circuit where the first circuit had a frequency of 100 kc and the second a frequency of approximately 400 kc (Fig. 4).

The periodically varying capacitances were diodes DGTs-25, DGTs-27 and D-809.

The use of a balanced circuit and of partial connection of the resonant circuits made it possible to get rid of the amplifier-circuit pumping voltage and of the harmonics of this voltage (which are brought about by the presence of a nonlinear capacitance in this circuit) without using additional filters.

Fig. 5 shows oscillograms of the voltages on the first and second circuits of a parametric amplifier; the time markers are spaced 100 microseconds apart.

Detuning produces damped oscillations in the transient. Their frequency is equal to the detuning if the input signal frequency is varied, and is less than that value (in our experiment by approximately a factor of 10) if the pumping generator or the second circuit are detuned.

The duration of the transient increases approximately in proportion to the voltage gain.

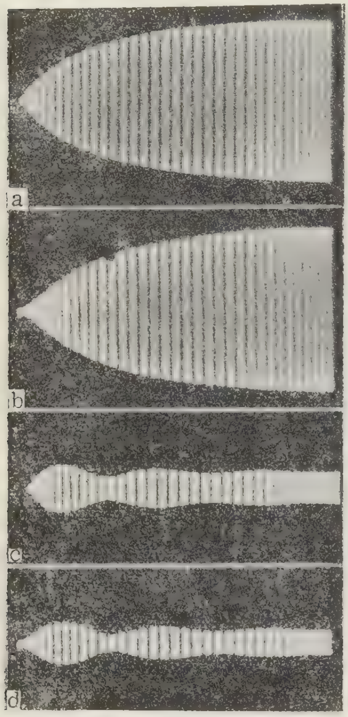


Fig. 5. Oscillogram of voltages on the first circuit (a), and on the second circuit (b) under optimal tuning and with the external-signal frequency detuning of 1 kc (c, d).

Physics Faculty, M. V. Lomonosov
Moscow State University

REFERENCES

1. J. M. Manly, H. E. Rowe, Proc. I. R. 1, 1956, 44, 7, 904.
2. H. E. Rowe, Proc. I. R. E., 1958, 46, 5, 850.
3. H. Heffner, G. Wade, J. Appl. Phys., 1958, 29, 9, 1321.
4. S. Bloom, K. Chang, RCA Rev., 1957, 18, December, 578.
5. M. D. Karasev, Uspekhi Fiz. Nauk, (Progress in Physical Sciences) 1959, 69, 2, 217.
6. N. N. Bogolyubov, Yu. V. Mitropol'skii Asymptotic Methods in the Theory of Nonlinear Oscillations, GIFML, 1958.
7. F. R. Gantmakher, Theory of Matrices, GTTI, 1954.

Received by editors June 1, 1960

SIMULATION AND CALCULATION OF TRAJECTORIES OF RELATIVIVISTIC CHARGED PARTICLES IN ELECTROSTATIC FIELDS

A.M. Strashkevich

As is well known, one of the clearest and most general methods of finding trajectories of nonrelativistic particles in an electrostatic field is the method of gravitational models. We have demonstrated [1, 2] that this model can be extended also to include the relativistic case, for which purpose the relief should be made not in accordance with the potential Φ (as in the nonrelativistic case), but in accordance with the so-called effective potential Φ^*. For Φ^* we have the following expression

$$\Phi^* = \Phi - \frac{1}{2}\frac{e\Phi^2}{m_0c^2}, \tag{1}$$

where e is the algebraic value of the charge, $c = 3 \times 10^{10}$ cm/sec, and m_0 is the rest mass. Although the trajectories of a relativistic particle in a field with potential Φ and of a nonrelativistic particle in a field with potential Φ^* (the motion of the latter is the one simulated in the proposed method) are the same, nevertheless, as indicated in [1, 2]) the variation of the velocities on these trajectories is different. Yet it is possible to obtain a simple relation between these velocities, a relation which can be used in the investigation of the motion of a relativistic particle by the method of effective potential, or by experimental methods based on the method of effective potential.

We denote by v_r and v_n the velocity of a relativistic particle (with mass m) and a nonrelativistic particle (with mass m_0) on identical trajectories (those of the relativistic particle in a field with potential Φ, and of the nonrelativistic particle in a field with potential Φ^*) and write accordingly the energy equations for these particles in the following fashion:

$$m_0c^2\left(\frac{1}{\sqrt{1-\frac{v_r^2}{c^2}}} - 1\right) + e\Phi = 0, \tag{2}$$

$$\frac{m_0v_n^2}{2} + e\Phi^* = 0. \tag{3}$$

We put, as is customary, $\left(1-\frac{v_\text{p}^2}{c^2}\right)^{-1/2} = \alpha$; from (2) we get

$$\alpha = 1 - \frac{e\Phi}{m_0c^2}. \tag{4}$$

In addition, we obtain from (2) the following expression for v_r:

$$v_\text{p} = \frac{c\sqrt{-e\Phi(2m_0c^2 - e\Phi)}}{m_0c^2 - e\Phi}. \tag{5}$$

From (2) and (1) it follows that

$$v_\text{H} = \frac{\sqrt{-e\Phi(2m_0c^2 - e\Phi)}}{m_0c}; \tag{6}$$

From (4), (5), and (6) we get

$$v_\text{H} : v_\text{p} = \alpha = m : m_0 \tag{7}$$

Consequently, the particles have the same value of momentum at the corresponding points of the identical trajectories considered here:

$$m_0v_\text{H} = mv_\text{p}. \tag{8}$$

Therefore, by photographing, for example, in intermittent light, a ball whose motion simulates the motion of a particle with mass m_0 and velocity v_n in a field with potential Φ^*, we can determine the law of variation of the momentum of a relativistic charge particle in a field with potential Φ; if it is desired to know the velocity v_r, we must divide the value of v_n by α, calculated from (4).

The same result can be obtained from the equations given in reference [3], in which the "time parameter" method is used to determine the trajectories of the relativistic particle. The gist of the method is that a certain parameter τ is introduced with the aid of the equation

$$\frac{dt}{d\tau}=\frac{m}{m_0}, \tag{9}$$

so that the relativistic equations of motion can be reduced to nonrelativistic ones (in form), but with the field intensity changed by a factor α. In fact, the equation of motion

$$\frac{d\left(m\frac{d\vec{l}}{dt}\right)}{dt}=e\vec{E} \tag{10}$$

changes upon substitution of t from (9) into

$$m_0\frac{\overrightarrow{d^2l}}{d\tau^2}=e\alpha\vec{E}. \tag{11}$$

From (9) we obtain our equation (7) if we recognize that τ is not simply a mathematical parameter introduced to transform the equations, but a parameter which has a well-defined physical meaning (something disregarded in [3]), namely that τ is the time of motion of a nonrelativistic particle in the field of the effective potential Φ^*. In fact, differentiating (1), we get

$$\operatorname{grad}\Phi^*=\alpha\operatorname{grad}\Phi; \tag{12}$$

Therefore, on going from an analysis of the motion of a particle with variable mass m in a field Φ to an analysis of the motion of a particle with constant mass m_0 in a field Φ^* (which indeed is the gist of the effective potential method, first proposed by Boguslavskiy [4]), we obtain equations identical with those obtained by the method of time parameter. Writing the main equation of this method (9) in the form

$$\frac{dl}{d\tau}:\frac{dl}{dt}=\alpha, \tag{13}$$

we arrive at our formula (7).

REFERENCES

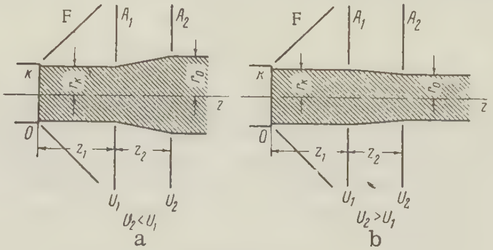
1. A. M. Strashkevich, Dokl. AN SSSR, 1956, 108, 3, 440.
2. A. M. Strashkevich, Electron Optics of Electrostatic Fields with No Axial Symmetry, Fizmatgiz, 1959.
3. A. Pignedoli, Z. angew. Math. und Mech., 1958, 38, 9/10, 343.
4. S. Boguslavskiy, Paths of Electrons in Electromagnetic Fields, Tr. Instituta fiziki, (Transactions of the Physics Institute), 1929.

Received by editors October 3, 1960

TWO-ANODE ELECTRON GUN FOR ELECTRONIC MICROWAVE DEVICES

Yu.V. Troitskiy

In the present communication we consider a two-anode electron gun, which produces an axially symmetrical electron beam emerging from the gun parallel to the axis and capable of being either greater or smaller than the cathode diameter. The geometry of the gun is shown schematically in Fig. 1. The focusing electrode F, located near the cathode K, has such a shape and such a potential, that the electrons leaving the cathode move parallel to the axis (as in the Pierce gun with parallel beam [1]). The beam formed in this region passes through apertures in the first and second anodes, which are at potentials U_1 and U_2.



a b

Fig. 1

In the case shown in Fig. 1a, the aperture in the first anode acts as a diverging lens, while the aperture of the second anode acts as a converging lens; in Fig. 1b the roles of the anodes are interchanged.

The sign of each anode lens depends on the electric field before and after the lens. Since it is assumed that the field is accelerating between the cathode and the first anode, and equal to zero between the first and second anodes, the sign of the lens is uniquely determined by the ratio of the anode potentials. By changing the anode potentials we can set the focal distance of the converging and diverging lenses in such a way, that the electron trajectories are parallel to the axis directly behind the second anode. This is a case favorable for further focusing of the beam by means of a longitudinal magnetic field; the magnetic field should then be applied to the beam at the point where the electrons have no radial velocities, i.e., immediately behind the plane of the second anode.

A calculation of the parameters of this gun was made under the simplifying assumption that the mutual repulsion between electrons can be neglected in the region between the cathode and the first anode the potential increases in the same manner as in a plane diode with current limited by space charge; in the region between the anodes, the potential varies linearly with distance; the potential of the space behind the second anode is equal to the potential of this anode. The beam was assumed to be paraxial, and the focal distances of the anode apertures were calculated by the formula of Davisson and Kalbick [formula (7.12) of reference [1]].

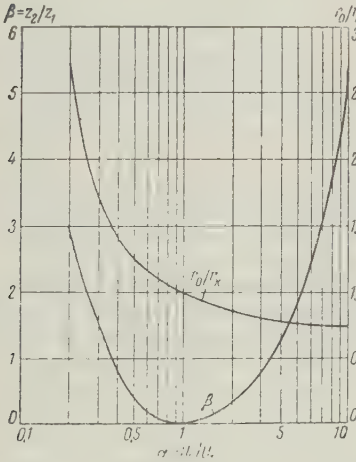


Fig. 2

The initial data for the calculation of the gun can be considered to be the following: the radius of the beam r_0 and the plane of the second anode, the beam current I_0, the potential U_2 on the second anode (which determines the velocity of the electrons after leaving the gun).

The design of the gun reduces to a determination of the distances z_1 and z_2 (Fig. 1), the radius of the cathode r_k, and the potential of the first anode U_1.

On the basis of the assumptions made and using formulas (7.12), (7.21) and (7.22) of [1], we can obtain by purely algebraic means the following formulas for the radius r_0 and for the slope of the trajectories $(dr/dz)_0$ of the boundary electrons of the beam after passing through the plane of the second anode:

$$r_0 = r_K\left[1 - \frac{\alpha - 1 - \frac{4}{3}\beta}{2(\alpha^{1/2}+1)}\right], \tag{1}$$

$$\left(\frac{dr}{dz}\right)_0 = -\frac{r_K}{8\alpha z_2}\left(3\alpha^{3/2} - 3\alpha - 3\alpha^{1/2} + 3 - 4\beta\alpha^{1/2} + \frac{4}{3}\beta\right). \tag{2}$$

Here we put $\alpha = U_2/U_1$; $\beta = z_2/z_1$.

We impose the condition $(dr/dz)_0 = 0$. This gives a relationship between α and β, valid when $z_2 \neq 0$:

$$\frac{z_2}{z_1} = \beta = \frac{9}{4}\,\frac{\alpha^{3/2} - \alpha - \alpha^{1/2} + 1}{3\alpha^{1/2} - 1}. \tag{3}$$

To Equations (1) and (3) we must add the connection between the current and the voltage in a plane diode, part of which is the region between the cathode and the first anode. On the basis of these three formulas we can determine from the specified values of I_0, r_0, and U_2 the unknown values z_1, z_2, U_1 and r_k. Since the number of unknowns is greater than the number of equations, one of these unknowns must be specified, for example r_k (which determines the density of the current on the cathode).

Fig. 2 shows the variation of β and r_0/r_k with α, obtained from formulas (1) and (3). For one and the same value of z_1/z_2, the gun can be both diverging ($\alpha < 1$, $r_0/r_k > 1$), or converging ($\alpha > 1$, $r_0/r_k < 1$). It should be noted that the assumptions made in the derivation of the formulas become meaningless near the point $\beta = 0$, when z_2 is small (on the order of the diameter of the apertures in the andoes) the region between anodes cannot be regarded as a combination of two thin lenses with a region of constant gradient.

A more exact calculation of the trajectories of the electrons between the anodes, with allowance for space charge, is given in reference [2]. The described gun construction is used in a low-noise traveling-wave tube for the 10 cm band with magnetically limited beam [3]. It should be noted that formula (35) in [3] differs from formula (3) of the present article, since it is derived without account of the focusing properties of the constant-field region between the anodes.

The formulas obtained were experimentally verified with a gun having $z_1 = 2.4$ mm, $z_2 = 1.75$ mm, and $r_k = 0.65$ mm. This verification was carried out with the aid of a special instrument, described in [4]. Measurements carried out at $U_1 = 200$ v have shown that the slope of the trajectories after leaving the second anode, at U_2 greater than U_1, vanishes for two values of α, 3.2 and 5. The first value is close to $\alpha = 2.9$, a quantity obtained by calculation with the aid of formula (3). The second figure, $\alpha = 5$, cannot be obtained from the formula and is apparently connected with the spreading of the beam under the influence of the space charge in the region between the anodes.

Calculation by means of formula (2) for $\alpha = 5$ shows that in the absence of space charge the beam should converge after leaving the gun. The mutual repulsion of the electrons in the beam in the space between the anodes reduces the angle of convergence and leads to the formation of a "neck" immediately behind the second anode.

The type of electron guns considered can be used in traveling wave and backward wave tubes and they are alike if it is necessary to obtain an electron beam parallel to the axis with small conductivity and with a diameter somewhat smaller or somewhat greater than the diameter of the cathode.

REFERENCES

1. J. R. Pierce, Theory and Design of Electron Beams, Van Nostrand, New York, 1954.
2. K. C. Ho, IRE Transactions 1955, ED-2, 3, 10.
3. F. N. H. Robinson, R. Compfner, Proc. I. R. E. 1951, 39, 8, 918.

4. Yu. V. Troitskiy, Pribory i tekhnika eksperimenta, Instruments and Experimental Techniques, 1957, 6, 113.

Received by editors September 29, 1960

DISTRIBUTION OF IMPURITIES IN THE REGION OF A WELDED CONTRACT ON GALLIUM ARSENIDE

S.N. Ivanov

Much attention is being paid at present to investigations of the properties of gallium-arsenide diodes with welded contact [1, 2]. In the present investigation we study the character of the distribution of impurities in the region of the rectifying contact, using a material having varying specific resistivity. The method used for the investigation involves the measurement of the dependence of the contact capacitance C on the reverse bias in the microwave range.

The diodes were made of gallium arsenide of n-type with resistivity $\rho = 0.006$ and 0.04 ohm-cm. The concentration of the majority carriers and the Hall mobility, obtained from measurements of the Hall constant and the specific resistivity, are listed in the table. To prepare the diodes, plates measuring 1.7 x 1.7 x 0.5 mm were cut. The etching was in a mixture of alkali and hydrogen peroxide [3]. A rectifying contact of small area was obtained in the following fashion. An alloy of 9% Zn + Sn was electrolytically deposited on a sharpened copper or silver wire (diameter of point about 5 microns), the point brought in contact with the crystal, and the welding effected by passage of a current pulse 0.25 — 0.5 amp for 10 milliseconds in the forward direction.

Table

ρ, ohm-cm	0.006	0.04
n, cm^{-3}	$5 \cdot 10^{17}$	$3 \cdot 10^{16}$
μ, cm^2/sec-v	1200	3000

An investigation of the static volt-ampere characteristics has shown that the dependence of the current I on the applied voltage U (in forward bias) has the form $I \sim e^{\alpha U}$, where $\alpha \simeq 30$ v^{-1}. To determine the law governing the distribution of impurities in the rectifying contact, we measured the contact capacitance C at frequencies $\omega = 2 \times 10^{10}$ and 6×10^{10} sec^{-1}, in accordance with the procedure of reference [4]. From the data obtained (see Figs. 1 and 2) it follows that the dependence of the capacitance C on the reverse bias U_{rev} has the form $C \sim (\varphi_c + U_{rev})^{-1/n}$, where $\varphi_c \simeq 1$ v. For diodes made of a material with $\varphi = 0.04$ ohm-cm. we have $n \simeq 3$; this dependence is characteristic of nearly linear distribution of the impurities in the region of the rectifying contact.

For diodes made of the material with $\varphi = 0.006$ ohm-cm, we have n = 2, showing an abrupt distribution of the impurities in the rectifying-contact region. The character of the dependence of the capacitance on the inverse bias was determined in our experiments only by the initial specific resistivity of the material and was independent of the material of the contact wire (copper or silver).

The presence of linear distribution of impurities in the region of the rectifying contact for diodes made of a material with $\rho = 0.04$ ohm-cm indicates that the welding is accompanied apparently by irreversible changes in the concentration of the impurities and defects in the small region near the contact. This can be connected either with impurity diffusion processes, or with formation of acceptor centers, owing to the large pressure of the point against the surface of the crystal and the high temperature.

To verify the last assumption, we carried out the following experiment. Samples of gallium arsenide with $\rho = 0.04$ ohm-cm, measuring 10 × 2 × 2 mm, were heated for one minute in a graphite holder with current, to a temperature of 700 — 800°C. A comparison of the measured Hall constant and resistivity before and after heating has shown that the resistivity and the Hall constant increase approximately by a factor of two. The Hall mobility does not change in this case. Analogous results of the influence of heat-treatment on resistivity of gallium arsenide are given in [5, 6]. Thus, heating of the material during the welding process can lead to an increase in the resistivity and to a change in the concentration of the impurities under the contact.

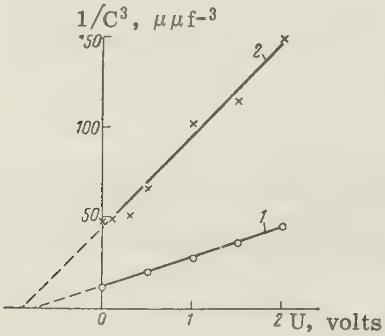


Fig. 1. Dependence of $1/C^3$ on the inverse bias for diodes made of material with $\rho = 0.04$ ohm-cm (1 — copper point, 2 — silver point)

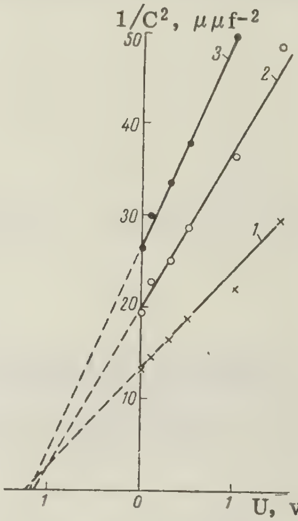


Fig. 2. Dependence of $1/C^2$ on the inverse bias for diodes made of material with $\rho = 0.006$ ohm-cm (1 and 2, points of silver, 3 — copper point)

When a material with $\rho = 0.006$ ohm-cm is used, the change in concentration of impurities and defects under the contact is small compared with the initial concentration of the impurities in the material, and a rectifying contact with abrupt distribution of impurities is obtained.

REFERENCES

1. D. A. Jenny, Proc. I. R. E., 1958, 46, 4, 717.
2. W. M. Sharpless, Bell. System Techn., J., 1959, 38, 1, 259.
3. D. N. Nasledov, A. Ya. Patrakova, B. V. Tsareknov, ZhTF (Journal of Technical Physics 1958, 28, 4, 779.
4. Yu. F. Sokolov, Radiotekhnika i elektronika, 1961, 6, 3, 399.
5. J. N. Allen, J. Electronics. and Control, 1959, 7, 3, 254.
6. J. J. Wysocki, J. Appl. Phys., 1960, 31, 9, 1686.

Received by editors February 4, 1961

THERMIONIC PROPERTIES OF CARBIDES OF HAFNIUM AND NIOBIUM AND OF BORIDES OF SCANDIUM AND THULIUM

V.S. Fomenko

To find new cathode materials, the emission and electric properties of various refractory compounds are presently under investigation, in particular, borides and carbides of rare and rare-earth metals [1 — 6].

In the present communication we describe investigations of the emission properties of scandium diboride, thulium hexaboride with small impurity of its tetraboride, and carbides of hafnium and niobium, on which there are no data in the literature.

Compounds for the investigation were obtained by procedures described in reference [7, 8].

Measurements were carried out by the procedure described in [5], using the method of Richardson curves in directly heated diodes with guard rings. The cathode was tantalum, tungsten, or molybdenum wire 0.3 mm in diameter, which served as the core. The wire was coated with a suspension of tantalum, tungsten, or molybdenum powder respectively, in a nitrobinder, to produce a rough cathode. The investigated substance was deposited on top of this in the form of a paste made of powder diluted in nitrobinder. The thickness of the finished cathode was 0.8 — 1.2 mm.

The cathode was calcined (thereby sintering the paste to the core) in an evacuated and outgassed tube. The calcining was at 700°K and about 10^{-5} mm Hg.

The cathode temperature was measured with a pyrometer through a special hole in the anode, a correction being made for the radiation of the material. The radiation coefficients of the compounds used in the investigation are listed in [9].

For thulium boride, an X-ray diffraction analysis was made of the initial substance and of the substance removed from the core of the cathode after measuring the thermionic emission characteristics. A comparison of these X-ray patterns has shown that no phase transformations take place after calcining of the material, disagreeing with the data given by the authors of reference [5] for analogous cases.

The results of the measurements are listed in the table.

Table

Substance	φ, ev	A, amp/cm² deg²	Temperature interval of measurement, degrees K	Density of emission current at temperature 400°K, amp/cm²
ScB_2	2.29 ± 0.08	$\sim 10^{-5}$	1100—1600	$5.0 \cdot 10^{-6}$
$TuB_6 (+ TuB_4)$	0.6 ± 0.9	$\sim 10^{-6}$	1100—1800	$1.69 \cdot 10^{-4}$
HfC	2.03 ± 0.2	$\sim 10^{-5}$	1100—1400	$7.5 \cdot 10^{-4}$
NbC	2.23 ± 0.04	$\sim 10^{-6}$	1200—1400	$9.5 \cdot 10^{-6}$

It should be noted that the same values of φ and A was obtained for cathodes deposited on different cores, this being due to the considerable thickness of the layer of the investigated material.

The low work function of the thulium boride can be attributed to the fact that the thermionic effect is masked here by β decay, and, consequently, the determination of the true work function of thulium boride is difficult, and perhaps impossible.

REFERENCES

1. D. Goldwater, R. Haddad, J. Appl. Phys., 1951, 22, 1, 70.
2. M. Lafferty, J. Appl. Phys., 1952, 22, 3, 299.
3. G. V. Samsonov, V. S. Neshpor, and G. A, Kudintseva, Radiotekhnika i elektronika, 1957, 2, 5, 631.
4. G. A. Kudintseva, B. M. Tsarev, V. A. Epel'baum. Collection: "Bor" (Boron), Trans. Conf. on Chemistry of Boron and Its Compounds, Goskhimizdat, 1958, p. 106.
5. G. A. Kudintseva, B. M. Tsarev. Radiotekhnika i elektronika, 1958, 3, 3, 428.
6. R. Pidd, G. Grover, D. Roehling, E. Salmi, J. Farr, N. Krikorian, W. Witteman, J. Appl. Phys., 1959, 30, 10, 1575.
7. G. V. Samsonov, Ya. S. Umanskiy, Solid Compounds of Refractory Metals, Metallurgizdat, 1957.
8. G. V. Samsonov, Yu. B. Paderno, Boridy redkozemel'nykh metallov (Borides of Rare-Earth Metals), Press of Acad. Sci. Ukr. S. S. R., 1960
9. T. I. Serebryakova, Yu. B. Paderno, V. S. Fomenko, Information Letter No. 245, Inst. of Metal Ceramics and Special Alloys, Acad. Sci. Ukr. S. S. R. 1960.

Institute of Metal Ceramics and
Special Alloys, Acad. Sci. Ukr. SSR

Received by editor September 7, 1960

CHRONICLE

JOINT SEMINAR ON CATHODE ELECTRONICS (NINETEENTH SESSION)

A.A. Yasnopol'skaya, R.P. Vikhlyayeva, T.M. Yekimenko

Eight papers were delivered to the nineteenth session of the Joint Seminar on Cathode Electronics, held on April 3, 1961 at the Institute of Radio Engineering and Electronics of the USSR Academy of Sciences.

N. D. Morgulis considered in his paper "The Problem of Direct Transformation of Heat into Electricity" the presently investigated new methods of effective conversion of heat into electricity, involving the use of principles of the physics of electronic phenomena in vacuum, gases, and solids. Such inversion methods can be either direct, such as thermoelectric or thermoelectronic, or semiindirect, such as magnetogasdynamic. The author considers the physical operating principles of such converters, the present status of their research and of the developed experimental models of such apparatus, as well as their possible applications to generation of electric power.

A paper "Certain Research on Cold Magnesium-Oxide Cathodes with Self-Maintaining Emission," written by a group of authors (N. Ya. Basalayeva, M. I. Yelinson, D. V. Zernov, T. M. Yekimenko, Ya. S. Savitskaya, A. A. Yasnopol'skaya), was delivered by M. I. Yelinson, The authors carried out several investigations of magnesium-oxide cathodes in order to ascertain the physical mechanism of their emission and to improve their working parameters; the technology of producing magnesium-oxide cathodes with the use of sputtering and cataphoresis methods was developed in detail; the thicknesses of the coatings were varied, and the granulometric composition, porosity, and roughness were determined.

It has been found that there exists an optimum amount of oxygen introduced into the magnesium-oxide layer. It has been established that the formation of an intermediate NiO layer between the core and the MgO layer is of no importance. It has been shown that the presence of oxygen is in principle essential, but its function is not merely to oxidize the remaining impurities. A reversible temperature dependence was established for the self-maintaining emission at the temperatures up to 400°C. The maximum of emission corresponds to a temperature of about 300°C. The presence of temperature-dependent irreversible changes in emission was also established. One of the causes of such changes is migration and removal of the oxygen. The tendency of the volt-ampere characteristics to display saturation is due to the effect of the temperature.

The time behavior of the emission was investigated. It was ascertained that the emission formation time exceeds 0.05 second and decreases with increasing emission current and with increasing starter-filament current.

The authors have established that self-maintaining emission depends on the magnetic field established parallel to the electric field in the magnesium-oxide layer. The minimum thickness of the layer at which it is possible to excite self-maintaining emission was determined.

The physical mechanism of emission is attributed to impact ionization in the grains of magnesium oxide, produced by the electrons accelerated in the pores, by hole conductivity through the grains of the magnesium oxide, due to the presence of oxygen which acts as an acceptor for the electrons formed during impact ionization, thus causing bending of the energy bands in the grains of the magnesium oxide and activation of the hole conductivity.

P. V. Timofeyev, N. D. Morgulis, and B. P. Nikonov participated in the discussions following the paper.

In a paper 'Investigation of the Interaction Between Molecules and the Surface of a Solid by the Molecular Beam Method' V. B. Leonas described a method of obtaining intense molecular beams and the results of research on their properties before and after their collide with the surface.

D. M. Mambetov delivered a paper "On one Form of Emission of Negatively Charged Particles from the Surface of a Dielectric Film on a Metal Substrate." In the investigation of α-radiation from uranium, the author and V. V. Cherdyntsev observed a large number of 'extraneous' pulses, decreasing to zero after a certain time. A detailed investigation has shown that the reason for the occurrence of the pulses is a thin layer of sodium bisulfate on a metallic substrate. The paper contains a description of the α-counting apparatus used to investigate the described phenomena. The author has advanced several hypotheses to explain the observed phenomena.

In the discussions, V. B. Sandomirskiy and G. N. Shuppe noted that the data observed are insufficient to explain the mechanism of emission.

A paper 'Interaction Between Barium Oxide and Tungsten in Vacuum' was delivered by B. P. Nikonov and O. M. Poltorak. The lecturers analyzed the possible reactions in a BaO–W system by comparing the pressure of oxygen in equilibrium with each of these substances in different states and at different temperatures, using the chemical-potential method. The analysis has shown that under no conditions can the pressure of the oxygen in equilibrium with the tungsten be less than in equilibrium with barium. The authors state that high vacuum is characterized by a state of dynamic equilibrium, in which $P_{Ba} = 2P_{O_2}$, where P_{Ba} is the pressure of the barium vapor, and P_{O_2} is the pressure of oxygen. Under these conditions the tungsten can be oxidized and lead to formation of nonstoichiometric oxides and their compounds with oxide of barium.

V. P. Kirsanov reported on the influence of molecular impurities on the work of the cathode of a zirconium lamp. Arc lamps with zirconium cathodes, being "point" sources of light, are lamps of the electrode-glow type. In the operation of these lamps it was observed that the brightness and the dimensions of the spot depend on the molecular gases present in the tube. The author has proposed an approximate method of calculating the size of the light spot.

A paper 'Concerning Bombardment of Barium-Oxide Films on Metallic Substrates at High Bombarding-Electron Energies' was delivered by P. V. Poshekhonov and Yu. V. Moskvichev. The authors made a mass-spectrometric analysis of the gases released during electron bombardment of molybdenum targets, coated with products of evaporation of the oxide cathode.

A. V. Druzhinin delivered a paper 'On the Nature of the Emitting Surface of Metal-Film Cathodes.' Electron-optical investigations of a L-type cathode, pressed tungstate cathode, and impregnated aluminate cathode has shown that their surface consists of an aggregate of highly emitting spots against a background of weakly-emitting surface with a work function of about 2.7 ev. The author proposes that the surface of the metal-film cathode of the normal activity is coated with a film of barium with a concentration less than optimal, and the observed bright spots are the consequence of the increase in concentration at the place where the barium emerges to the surface or a result of the "showing through" of the productive face under the layer of the barium.

A. A. Yasnopol'skaya, R. P. Vikhlyayeva, T. M. Ekimenko

RADIO ENGINEERING AND ELECTRONIC PHYSICS

Institute of Radio Engineering and Electronic Physics,
Academy of Sciences of the USSR

EDITORIAL BOARD

Editor-in-Chief: V.A. Kotel'nikov
Associate Editors: D.V. Zernov, Yu.B. Kobzarev

A.I. Berg
B.A. Vvedenskiy
I.S. Gonorovskiy
V.L. Granovskiy
L.A. Zhekulin
N.D. Devyatkov
L.N. Dobretsov
A.N. Kazantsev
S.G. Kalashnikov
P.L. Kapitsa
V.V. Migulin
A.L. Mikaelyan
A.A. Pistol'kors
A.M. Prokhorov
S.M. Rytov
V.I. Siforov
Ya.N. Fel'd
S.E. Khaykin
B.M. Tsarev

Scientific Secretary of Editorial Board: G.A. Bernashevskiy

The English Edition of Radio Engineering and Electronic Physics is mailed to subscribers within 18 weeks after the publication of the original Russian issue.

Russian electronic journals published by the
American Institute of Electrical Engineers
Translated by Royer and Roger, Inc.

	Subscription rates			
	Individuals		Libraries	
	$	£	$	£
Radio Engineering and Electronic Physics	28.50	10	57.00	20
Radio Engineering	14.25	5	28.50	10
Telecommunications	14.25	5	28.50	10

Royer and Roger translates and produces
the following Russian scientific journals:

Biophysics
Entomological Review
Geochemistry
Geodesy and Cartography
Izvestiya, Academy of Sciences of the USSR, Geologic Series
Pavlov Journal of Higher Nervous Activity
Problems of Oncology
Radio Engineering
Radio Engineering and Electronic Physics
Refractories
Sechenov Physiological Journal of the USSR
Soil Science
Telecommunications

Comments and inquiries regarding *Radio Engineering and Electronic Physics* and other translation journals should be sent to:

International Division
Royer and Roger, Inc.

1000 Vermont Avenue, N.W.
Washington 5, D.C.

41 East 28th Street
New York 16, New York

RADIO ENGINEERING AND ELECTRONIC PHYSICS

РАДИОТЕХНИКА И ЭЛЕКТРОНИКА

NUMBER 8 1961

CONTENTS

N.A. Armand and B.A. Vvedenskiy, Allowance for Antenna Directivity Pattern in the Calculation of Radio Wave Diffraction around the Earth 1083
P.Ye. Kubasov, Investigation of Space-Time Correlations of the Field Level of Centimeter Waves Propagating in a Turbulent Atmosphere 1091
Ya.D. Shirman, Statistical Analysis of Optimal Resolution 1098
G.D. Malyuzhinets and L.A. Vaynshteyn, Transverse Diffusion in Diffraction by an Impedance Cylinder of Large Radius. Part I. Parabolic Equation in Ray Coordinates 1106
M.D. Khaskind, Excitation of Waves on an Impedance Plane 1117
B.Ye. Kinber, Diffraction of Electromagnetic Waves by Concave Surface of Circular Cylinder 1130
R.B. Vaganov and V.V. Meriakri, Suppression of Resonance Phenomena in Multimode Waveguides 1139
P.P. Pavlov, Electromagnetic Field and Distribution of Current along an Infinitely Long Insulated Wire in a Conducting Medium. 1148
S.I. Viglin, Transient Characteristics of Coaxial System with Periodic Structure . . . 1161
A.B. Lazarev and G.F. Filimonov, Static Characteristics of a Plane Magnetron. 1169
G.A. Postnov, Possibility of Existence of Slow Waves in an Isotropic Plasma Waveguide 1176
G.P. Apushinskiy, Concerning the Sensitivity of Radiometers 1179
Z.S. Gribnikov, B.N. Kompan and L.V. Svyatogor, Investigation of Field Effect Transistors 1181
M.I. Yelinson, F.F. Dobryakova, et al., On the Theory of Field Emission and Thermionic Field Emission of Metals and Semiconductors. 1191
Sh.M. Kogan, Thermal Radiation Produced by Hot Carriers in Semiconductors 1201
G.B. Der-Shvarts, Yu.M. Kushnir, et al., Certain Problems of Reflex Electron Microscopy, ... 1204
V.N. Vertsner, V.G. Nikiforova, et al., The EM-6 Light-Electron Microscope. . . . 1209
R.Ya. Berlaga, P.P. Konorov, M.I. Rudenok, Electron-Microscopic Investigations of Germanium Surface 1214
A.N. Pilyankevich, Calculation of the Characteristics of Scattering of Medium-Energy Electrons on Free Atoms 1217
P.A. Stoyanov, N.M. Grishina, Elimination of Astigmatism of the Intermediate Lens of an Electron Microscope with the Aid of a Stigmator During Microdiffraction . . 1220
P.A. Stoyanov, Alignment of an Electron Microscope by Means of Electromagnetic Fields 1225
Yu.V. Vorob'yev, Analysis of Aberrations Due to Defects in Manufacture and in Alignment of Electron Microscopes. 1229

BRIEF COMMUNICATIONS

A.V. Prosin, On the Calculation of the Bandwidth in UHF Tropospheric Scatter Propagation 1234
B.M. Minkovich, Yu.I. Davidchevskiy, On the Synthesis of an Antenna with Circular Aperture 1237
Yu.A. Il'inskiy, M.D. Karasev, Transients in a Two-Circuit Parametric Amplifier 1239
A.M. Strashkevich, Simulation and Calculation of Trajectories of Relativistic Charged Particles in Electrostatic Fields. 1243
Yu.V. Troitskiy, Two-Anode Electron Gun for Electronic Microwave Devices. 1245
S.N. Ivanov, Distribution of Impurities in the Region of a Welded Contract on Gallium Arsenide 1247
V.S. Fomenko, Thermionic Properties of Carbides of Hafnium and Niobium and of Borides of Scandium and Thulium 1249

CHRONICLE

A.A. Yasnopol'skaya, et al., Joint Seminar on Cathode Electronics (Nineteenth Session) 1251